Interferometric-Spatial-Phase Imaging for Sub-Nanometer Three-Dimensional Positioning

by

Euclid Eberle Moon

S.B., Physics, Massachusetts Institute of Technology
S.M., EECS, Massachusetts Institute of Technology

Submitted to the Department of Electrical Engineering
and Computer Science in partial fulfillment of the requirements
for the degree of

DOCTOR OF PHILOSOPHY IN THE DEPARTMENT OF ELECTRICAL ENGINEERING AND COMPUTER SCIENCE
at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY
September 2004

Signature of Author.....

Department of EECS September 7, 2004

Certified by.....
Henry I. Smith
Keithley Professor of Electrical Engineering
Thesis Supervisor

Accepted by.....
Arthur Smith
Chairman, Graduate Thesis Committee

Interferometric-Spatial-Phase Imaging for Sub-Nanometer Three-Dimensional Positioning

by

Euclid Eberle Moon

Submitted to the Department of Electrical Engineering and Computer Science
on September 7, 2004 in Partial Fulfillment of the Requirements for the Degree of
Doctor of Philosophy in Electrical Engineering and Computer Science

Abstract

Current alignment technology is incapable of satisfying the needs of imminent generations of lithography. This dissertation delineates a novel method of alignment and three-dimensional position metrology that is compatible with many forms of proximity lithography. The method is called Interferometric-Spatial-Phase Imaging (ISPI), and is based on encoding three-dimensional position information in the spatial phase and frequency of interference fringes, viewed with specialized oblique-incidence, dark-field optical microscopes. Alignment detectivity is <0.5 nm, and detection range is >500 μm. Unlike amplitude-based interferometers, this spatial-phase-encoding interferometer achieves high alignment detectivity without sensitivity to variations in wavelength, gap and other factors, such as resist layers and changes in the index of refraction in the beampath. Several novel gap detection methods are introduced, with gap detectivity <1 nm, measured over gaps between <1 μm and >500 μm. Gap is confirmed with exposure of patterns in resist, taking advantage of near-field interference in a novel Chirped Talbot Effect. Alignment and pattern overlay are confirmed in experiments combining x-ray exposures with continuous ISPI position feedback. Dynamic overlay of patterns in resist is demonstrated to be 2.7 nm, with a clear path for further improvement. Gate structures in a double-gate MOSFET are dynamically aligned to 2.5 nm.

Thesis Supervisor: Prof. Henry I. Smith
Title: Keithly Professor of Electrical Engineering and Computer Science

Acknowledgements

My most profound gratitude goes to Prof. Hank Smith, who has been an invaluable mentor, an inexhaustible source of enthusiasm, and a paragon of incisive thought and action. Hank's edifying methods of scientific investigation, inspirational presentation style, and clarity of writing forever will remain my ideals. It goes without saying that this work would have been impossible without his enduring intellectual, emotional, and material support.

Many thanks to go to my friends and coworkers at the NSL, SNL, and other labs:

- Mark Schattenburg, for intense discussions on nano-precision and macro-insights.
- Rajesh Menon, for his good humor, intense curiosity, an introduction to psytrance, and raves in grungy New York alleys and the California wilderness.
- Nitin Gupta, for introducing me to mountain biking at Killington and the Blue Hills, and most importantly, for Dream Theater.
- Paul Herz, for telling it like it is and the odyssey in Utah.
- Tymon Barwicz, for his visual perspicacity and vivid images of Greece.
- Mark Finlayson, for his crackling wit and dance moves.
- Joy Cheng, for injecting brightness and energy, and making 6.730 even more enjoyable.
- Alberto Moel, for opera renditions at 3AM in the cleanroom.
- Mitch Meinhold, for NanoWriter and making 72-hour exposures a great experience.
- Feng Zhang, for being a (literally) supportive workout buddy.
- Mike Walsh, for his ready advice in my deepest FDTD quandaries.
- Minghao Qi, for teaching me Chinese chess (and enabling me to print this thesis).
- Pat Everett, for his attention to detail, hard questions, and design of good experiments.
- David Breslau, for his impressive mechanical insight and virtuostic craftsmanship.
- Mark Mondol, for his warmth, breadth of experience, introduction to some of the best Indian and Tibetan restaurants, and memorable Christmas parties.
- Prof. Philip Morrison, for letting me follow my own compass and set off into this field in by bachelor's thesis.

Dedication

This dissertation is dedicated to my parents, without whom none of this would have been possible.

Contents

List of Figures

List of Tables

'Increasing knowledge brings with it increasing power, and great as are the triumphs of the present century, we may well believe that they are but a foretaste of what discovery and invention have yet in store for mankind. The work may be hard, and the discipline severe; but the interest never fails, and great is the privilege of achievement.' - Lord Raleigh

1 Introduction

1.1 Motivation and Objectives

Alignment and gap control, and position metrology in general, are pervasive requirements in nanofabrication. Throughout the history of micro- and nano-fabrication, fine three-dimensional positioning has always been a key enabling technology. Indeed, alignment and metrology were cornerstones of each of the previous industrial revolutions over the past 200 years[1], and are likely to continue to be indispensable partners with any foreseeable fabrication technology.

[1] Metrology was integrated with machining via Henry Maudslay's invention of a micrometer capable of measuring to 1/10,000 of an inch (2.5 μm) in 1805 [1.1], permitting control of milling machines for manufacture of interchangeable parts.

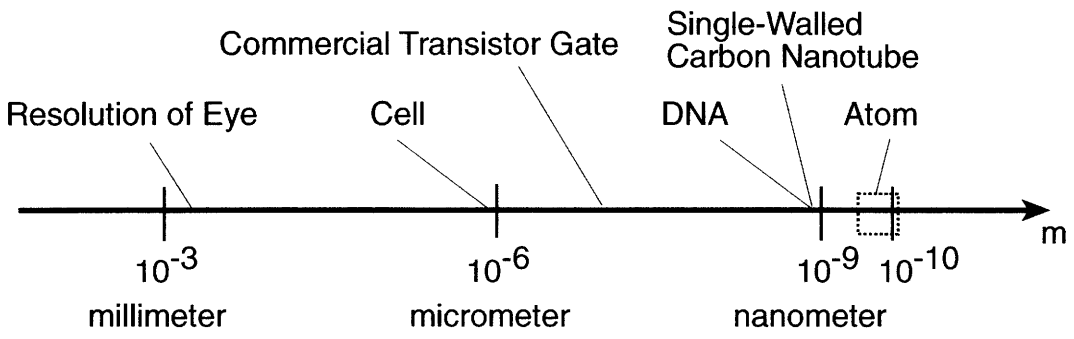


Fig. 1.1. Scale of microscopic dimensions.

Currently, the semiconductor industry produces transistor gates with 90 nm widths, and requires alignment and pattern overlay of 30 nm. The importance of new alignment methods is indicated by the industry-wide International Technology Roadmap for Semiconductors (ITRS) 2003 [1.2], which states that there is "no known solution" to alignment beyond the year 2007 when overlay is specified to be 23 nm. Further in the future, as predicted by the ITRS, alignment accuracy demands are expected to increase to <7 nm. This viewpoint is reflected by sources in the nascent nanofabrication industry [1.3], who say "The challenge of manufacturing standards remains one of the most formidable obstacles to widespread commercialization: Nanomanufacturing needs accuracy down to 5 or 10 nanometers, ***something nobody knows how to do***. We need much, much better registration ..., and it's going to be a real challenge."

In the author's viewpoint, it is virtually certain that continued progress in nanofabrication will depend upon exquisitely fine position control, including control with sub-nanometer accuracy. When the end of the semiconductor roadmap is reached, there will be numerous applications of alignment that involve the interface between the molecular and lithographic regimes. For instance, in templated self-assembly, in which molecules form regular arrays between template walls [1.4], the spacing of templates by fractions of the molecular diameter determine the packing arrangement and exact number of molecules that can self-assemble in the interstitial regions.

This dissertation will respond to the expansive and enduring nano-alignment challenge with a unique three-dimensional position metrology technique that provides *sub-nanometer* detectivity, exceptional dynamic range, high signal-to-noise ratio, immunity to numerous perturbations and deleterious conditions, and applicability to any foreseeable form of proximity lithography.

1.2 Background of Microprocessor Fabrication

This section will outline the development of the microprocessor, with lithography, the planar process, and alignment as three essential enabling technologies. Certainly there are a plethora of other applications of alignment, but we choose microprocessor fabrication as a vehicle to illustrate a broadly useful application to which some form of alignment has been indispensable throughout its history.

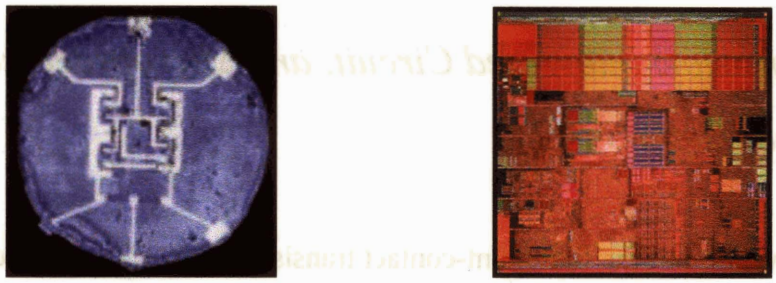
1.2.1 Transistor, Integrated Circuit, and Microprocessor Development

The invention of the first point-contact transistor (Fig. 1.2) on December 16, 1947 by Bardeen, Brattain and Shockley (patent #2,524,035) initiated the unprecedented and wide-ranging revolution in technological capability that is the foundation for our information-rich society.

(a) (b)

Fig. 1.2. (a) Bardeen, Shockley and Brattain (respectively). (b) The first transistor.

The transistor was made practical by the first semiconductor junction transistor invented in 1951 by Shockley, and scalable, by the inventions of the first integrated circuit in 1958 by Kilby [1.3] on germanium (patent #3,138,743), and the first integrated circuit on silicon by Noyce in 1959 (patent #2,981,877). The first commercially available integrated circuit followed in 1961, produced by Fairchild Semiconductor.



(a) (b)

Fig. 1.3. (a) First integrated circuit, with two transistors. Alignment registration was ~100,000 nm. (b) Intel Prescott chip: 125 million transistors with 90 nm features. Alignment requirement is 30 nm.

As shown in Fig. 1.3, feature sizes on the transistors in early integrated circuits were on the order of 100s of μm, and the overall size was ~1 cm (for two transistors).

Early integrated circuits were a tremendous improvement over vacuum tubes[2], in terms of weight, size, cost, and reliability, and over previous fabrication technology (i.e., hand-assembly and soldering), in terms of practicality, yield, and, most importantly, scalability. Prior to the integrated circuit, transistors required manual soldering of components, and the variation of the components resulted in inconsistent operation. Integrating all the components on a single (silicon) substrate meant that the components were matched to each other, and no manual intervention was required, increasing the rate of device production, as well as lowering cost. More transistors could be added to the same substrate at minimal increases in cost or process complexity, leading to the exponential growth in the number of transistors per chip, known as Moore's Law (Fig. 1.4).

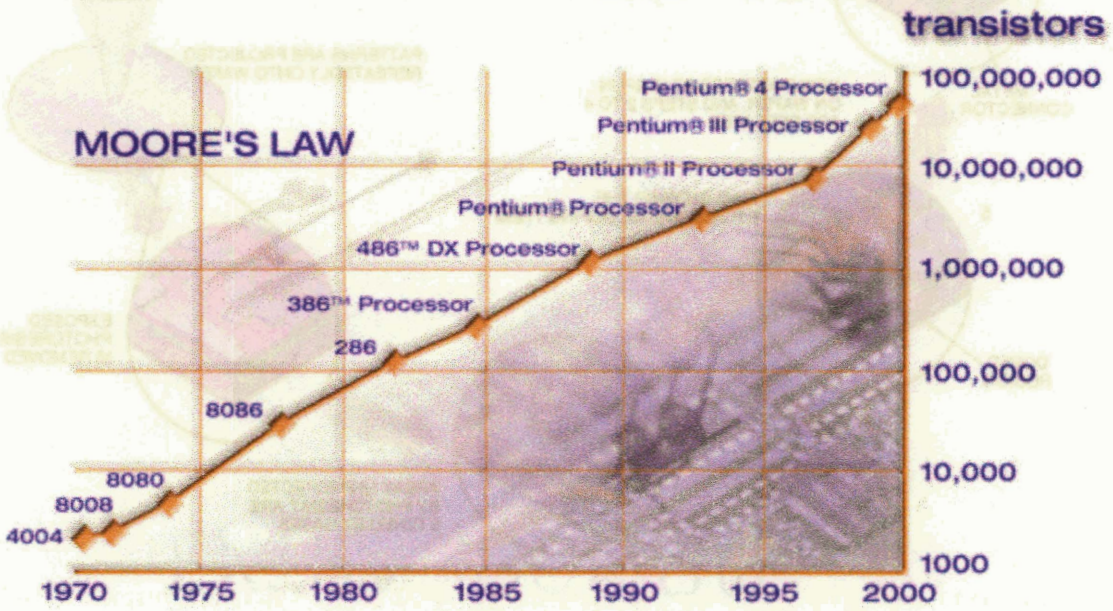


Fig. 1.4. Semi-log plot of the number of transistors per microprocessor versus time.

[2] Transistors made a critical difference in the implementation of digital computers. In the late 1960's IBM made the first digital computers using transistors. Up to that time, digital computers were made with vacuum tubes, which were comparatively large, heavy, and drew far more power. For instance, an ENIAC digital computer, built in 1945, had 19,000 vacuum tubes and drew 200 kW. A modern personal computer contains ~100 million transistors and draws ~200 W.

1.2.2 *Lithography and the Planar Process*

The technologies that made microprocessors a reality were lithography and the planar process, as illustrated in Fig. 1.5.

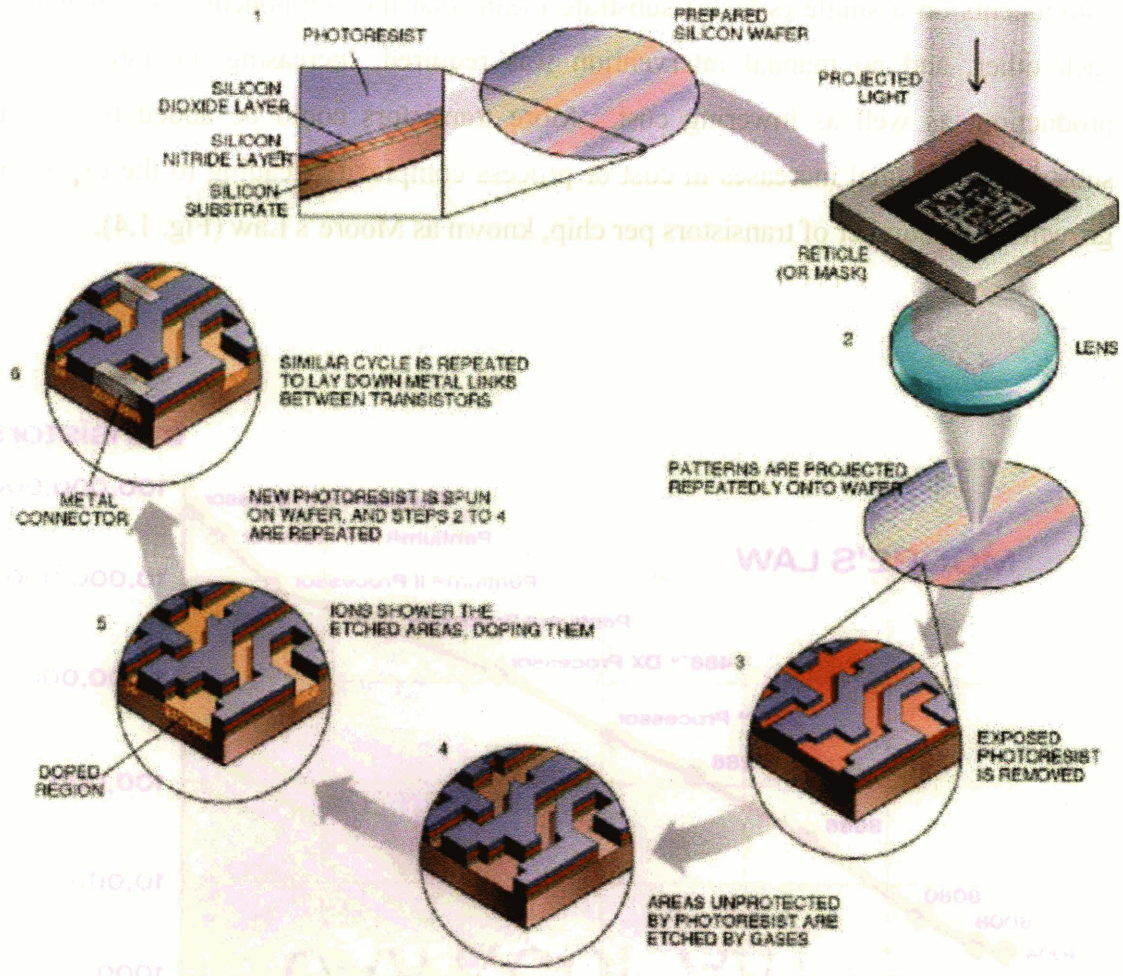


Fig. 1.5. Schematic of lithographic processes. Alignment and proper overlay is required upon each of the 20 or more lithographic cycles required to make a circuit.

In lithography, a pattern is transferred into a photosensitive film (resist) by means of chemical changes in the resist resulting from exposure to light, electrons, or other particles. The particle flux may be in the form of a narrow beam scanned over the

surface, or in parallel, from a fixed pattern defined by a mask. In a positive-tone resist, the particle-induced changes to its chemical structure make it susceptible to removal with a chemical agent (a developer solution), while unexposed regions are unaffected, creating a relief pattern. The resist pattern protects underlying regions of the substrate, while further processing steps are performed on the unprotected regions. The pattern is transferred into the substrate by wet or dry etching, or layers of metal or insulator deposited in the interstitial regions. Alternatively, the resist pattern can selectively block penetration of ions – in the unblocked regions, injected ions dope the silicon to alter its semiconductive properties. When the resist is removed, a fixed layer of semiconductor, metal, or insulator remains, in well-defined lateral and vertical dimensions. A new resist layer is spun on, and the cycle is repeated layer by layer, with various patterns in each layer, to form the complete circuit.

The lithographic process is repeated >20 times (21 times, in the case of a Pentium 4) to define the various parts of a device, such as a gate (or gates), source, drain, contact pad or interconnect. Misalignment by >1/3rd of the gate length will cause a variety of unacceptable effects.

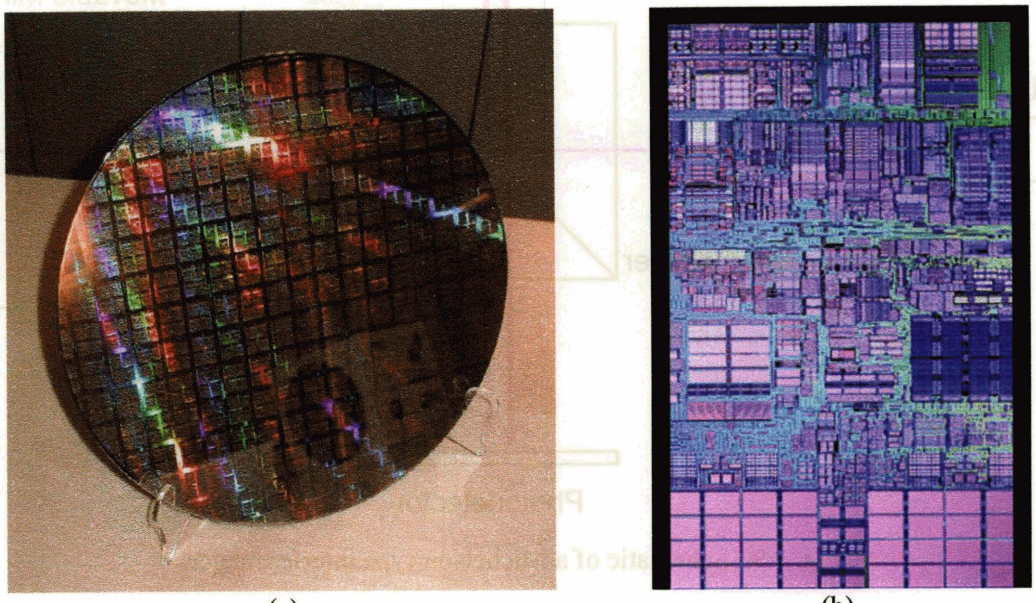
The rapid growth of the number of transistors on a chip depends upon the ability to fabricate as many devices as desired in parallel, within the lithographic resolution limits. The type of lithography is critical, and frequently is the limiting factor to scaling up to more transistors in smaller area.

On a historical note, the first known instance of lithography is attributed to the Hohokam Indians of southern Arizona, around 1000 B.C [1.5, 1.6]. The Hohokam Indians imported seashells from the California coastline, masked regions with patterns of mesquite tar pitch, and etched the uncovered regions of the shell with fermented juice from the fruit of the saguaro cactus. Figure 1.6 shows an example of a Hohokam shell.

Fig. 1.6. Seashell patterned by the Hohokam process (circa 1000 B.C.). Patterns were masked by tar pitch, and etched by fermented juice (weak acetic acid) from the fruit of the saguaro cactus.

In contrast to the etched seashell, an example of contemporary state-of-the-art fabrication facilities and their products are shown in Figs 1.7 and 1.8.

Fig. 1.7. IBM East Fishkill semiconductor fabrication facility.



(a) (b)

Fig. 1.8. (a) 300 mm-diameter PowerPC G5 wafer produced in IBM East Fishkill fabrication facility. The wafer contains several hundred chips. (b) A G5 chip layout, containing 58 million transistors. Minimum feature sizes are 90 nm, and alignment <32 nm for each transistor gate. Lithography is done with the 193 nm wavelength.

1.3 Background of Interferometric Position Detection

1.3.1 Homodyne Interferometer

The first step in performing fine alignment is to detect position to less than the wavelength of light. The most widely used tool for fine position metrology is the laser interferometer. The origin of interferometry dates back more than 200 years. In 1803, Young performed the first interference experiments[3] with a double-slit, wavefront-splitting interferometer. Michelson[4] developed the first amplitude-splitting interferometer in the 1880's, providing a comparison between two path lengths, as shown in Fig. 1.9.

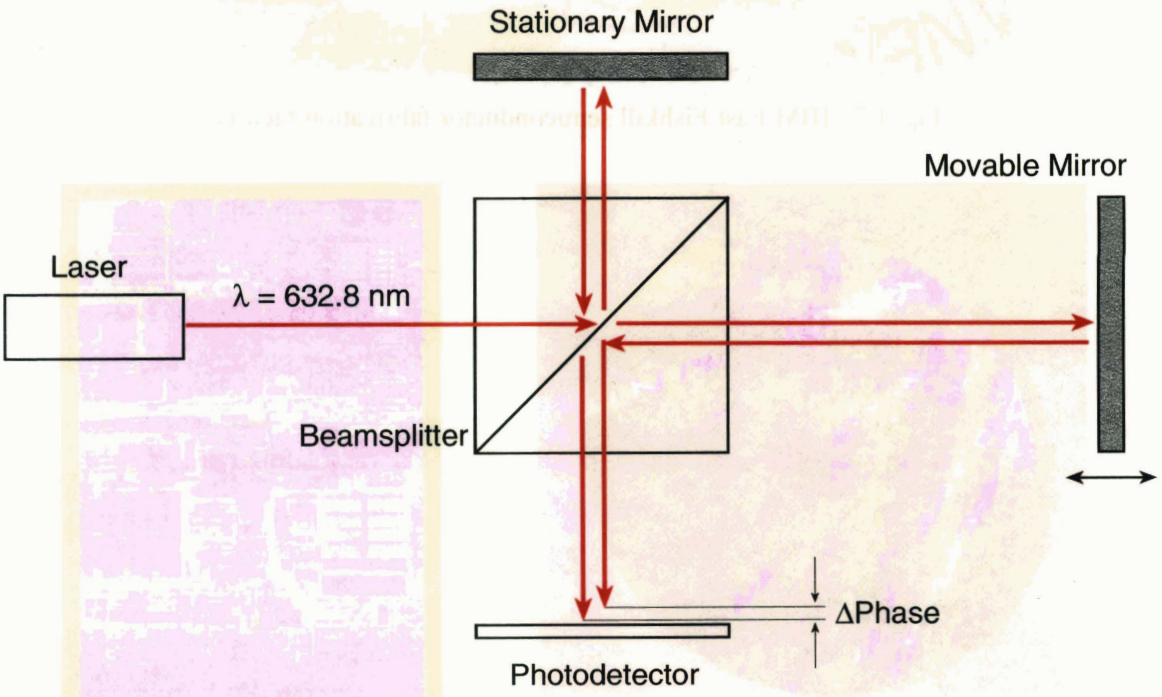


Fig. 1.9. Schematic of a Michelson-type interferometer.

[3] Young's double-slit interferometer provided the best evidence at the time in support of a wave theory of light, and provided the pivotal counter-argument to Newton's corpuscular theory.
[4] Nobel Prize winner, 1907

In this interferometer a collimated, monochromatic beam is divided by a beamsplitter into two equal parts, with one beam reflecting from a fixed mirror, and the other beam reflecting from a mirror attached to a moving platform. The two reflected beams recombine at the beamsplitter, and the interferometric intensity is detected by a photodetector. Moving one mirror changes the relative phase between the interfering beams, resulting in an intensity variation at the detector, e.g., a path length change of $\lambda/2$ corresponds to one cycle in the intensity of the interfering beams.

The two collinear plane waves are expressed as

$$E_1 = A_1 e^{i(\omega t + \phi_1)}$$
$$E_2 = A_2 e^{i(\omega t + \phi_2)} \tag{1.1}$$

where ϕ_1 and ϕ_2 are their phases. Their superposition results in an intensity I, where

$$I = E_1 E_2^* = I_1 + I_2 + 2\sqrt{I_1 I_2}\cos(\phi_1 - \phi_2) \tag{1.2}$$

with an interference term that varies sinusoidally; the argument being the phase difference (ϕ_1 - ϕ_2) between the two waves. The phase difference depends on the additional path length traveled by the beam to the movable mirror. Hence, the displacement of the mirror changes the intensity of the interfering beams in a cyclical fashion.

Commercial interferometers[5] based on a single frequency interference (homodyne) were introduced in 1964, soon after the invention of the helium-neon laser[6], and could detect displacements as small as $\lambda/8 = 79$ nm.

[5] Airborne Instruments Labs, Inc.
[6] Ali Javan, 1960.

Problems with the Michelson interferometer

- The photodetector measures only the intensity of light at a point, so any variation in the amplitude of the signal due, for example to laser power fluctuations, ambient light pickup, or electronic noise, will be interpreted as a displacement. Over long time scales, drift in the alignment of the optical components will also vary intensity, and bias displacement measurements.

- There is an inherent phase ambiguity: every motion of the mirror through $\lambda/2$ results in the same intensity, therefore, the total displacement is determined by counting cycles in the intensity oscillations. A change in source intensity or detector response will be mistaken as a displacement unless a second detector and additional interferometer paths are employed in a quadrature scheme.

- The direction of motion cannot be sensed without using a quadrature scheme.

- Changes in density of air in the beampath vary the index of refraction, causing optical path length variations in one arm of the interferometer. This causes intensity changes, even though the mirrors are stationary.

1.3.2 Heterodyne Interferometer

Many of the shortcomings of the homodyne interferometer were overcome by detecting displacement using the phase difference of two frequencies: i.e., in a heterodyne interferometer. In the heterodyne interferometer, the phase difference is due to temporal beats between the two frequencies, in distinction to the phase differences arising from the path length differences in the Michelson interferometer. In the heterodyne scheme, orthogonally polarized light of frequencies f_1 and f_2 (resulting from

Zeeman splitting or, more commonly, acousto-optic modulation[7]) is emitted from a dual-frequency HeNe laser, Fig. 1.10.

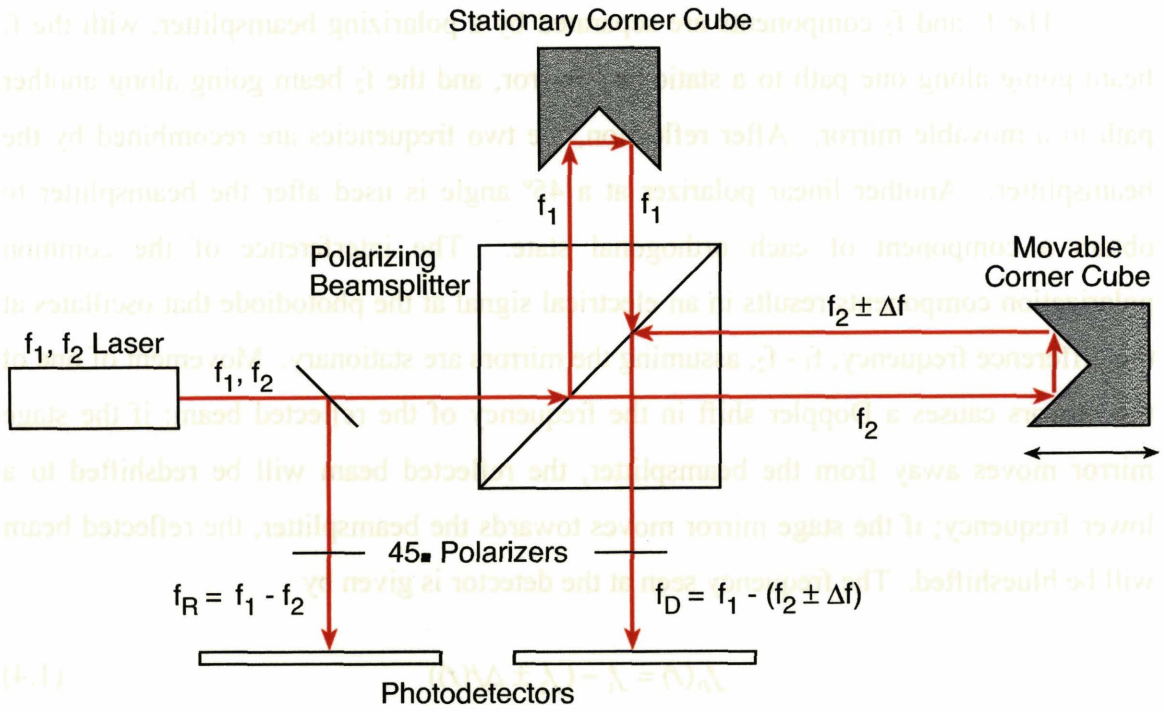


Fig. 1.10. Schematic of a heterodyne interferometer. A dual-frequency laser produces two frequencies in orthogonally-polarized states. Displacement of a movable corner cube is encoded in the Doppler frequency shift, compared to a stationary corner cube.

The laser output is split initially into two parts by a non-polarizing beamsplitter, one part for reference, and the other for measurement. The reference beam passes through a 45° polarizer, to obtain interference between components of each of the orthogonally polarized states. A reference frequency f_R is detected by intensity oscillations on a photodetector, and is given by

$$f_R = f_1 - f_2 \tag{1.3}$$

[7] The acousto-optic method uses a frequency shifter, such as a Bragg cell, to produce the frequency difference. This technique yields a frequency split (20MHz) that is much greater than that of the Zeeman technique. The split also remains constant because the Bragg cell is driven by a stable quartz oscillator.

which is independent of time (within the bounds provided by the frequency stability of the laser).

The f_1 and f_2 components are separated by a polarizing beamsplitter, with the f_1 beam going along one path to a stationary mirror, and the f_2 beam going along another path to a movable mirror. After reflection, the two frequencies are recombined by the beamsplitter. Another linear polarizer at a 45° angle is used after the beamsplitter to obtain a component of each orthogonal state. The interference of the common polarization components results in an electrical signal at the photodiode that oscillates at the difference frequency, $f_1 - f_2$, assuming the mirrors are stationary. Movement of one of the mirrors causes a Doppler shift in the frequency of the reflected beam: if the stage mirror moves away from the beamsplitter, the reflected beam will be redshifted to a lower frequency; if the stage mirror moves towards the beamsplitter, the reflected beam will be blueshifted. The frequency seen at the detector is given by

$$f_D(t) = f_1 - (f_2 \pm \Delta f(t)) \tag{1.4}$$

where $\Delta f(t)$ is the amount of the Doppler shift, and is given by

$$\Delta f(t) = 2nf_2 v(t) \tag{1.5}$$

where n is the index of refraction of the medium in the beam path and v is the stage velocity.

Electronics measure both f_D and f_R, and determine the displacement of the stage mirror by converting the phase

$$\phi(t) = 2\pi \int_0^t \Delta f(\tau) d\tau = 2\pi \int_0^t \frac{2nv(\tau)}{\lambda} d\tau = \frac{4\pi n}{\lambda} \int_0^t v(\tau) d\tau \tag{1.6}$$

into displacement by the relation

$$D(t) = x(t) - x(0) = \int_0^t v(\tau) d\tau = \frac{\lambda}{4\pi n} \phi(t) \tag{1.7}$$

When first introduced in the Hewlett-Packard 5525A [1.7] in 1970, heterodyne interferometers improved detection to $\lambda/64$ (633 nm/64 $\approx$ 10 nm). Further improvements in electronics allow detection to $\lambda/2048$ (633 nm/2048 = 0.3 nm). Multiple passes of the laser beam lengthen the overall beampath by folding the beampath onto itself in each arm, allowing an increase in sensitivity to ~0.15 nm (for a 4-pass configuration).

The heterodyne interferometer combines both high resolution (0.15 nm) and long range (>10 m), and can measure high velocities (up to 4.2 m/s).

An important characteristic of interferometry is that only the displacement is measured, not the absolute position. Therefore, the initial distance to the movable mirror or corner cube is not measured, only the change in relative position between the reflectors. To find the absolute position, cycles from a reference point must be counted. An advantage of heterodyne interferometers is that they can measure the direction of motion without any additional apparatus, such as is needed in the homodyne quadrature scheme.

The primary advantage of the heterodyne interferometer is that by encoding position in phase, instead of amplitude, changes in background intensity (e.g., turning on a light in the room) would not cause a change in the interferometer signal that could be interpreted as a displacement.

Problems with heterodyne interferometers

Frequency drift of the laser:

A frequency drift can be interpreted as a displacement. Temperature variations can cause drift in the length of the laser cavity, and the resonant frequency. A heating coil around the laser tube is commonly used to stabilize the laser temperature.

Index of refraction:

Heterodyne measurements assume a consistent index of refraction in both the reference path and the Doppler-shifted paths, which is not always a valid assumption due to turbulence and density variations in the air, variations in humidity, or mixing of air with other gases such as helium. These variations are several orders of magnitude greater than errors arising from frequency drift.

1.4 Obstacles to Nano-Alignment

Laser interferometers demonstrate feasibility of sub-nanometer position sensitivity, but are insufficient to achieve robust, nanometer-level pattern overlay. We can summarize the obstacles with three essential points:

1. Interferometers are subject to environmental disturbances.
2. Interferometers measure the position of mirrors at the edges of stages, not mask-substrate position.
3. Thermal and mechanical instability on the nanometer scale cause deviations from the position of a stage mirror and the position of a feature on a mask.

Environmental controls and low-thermal-expansion-coefficient materials attempt to stabilize dimensional variations, however, without direct knowledge of the relative position of a mask and substrate, nanometer-level overlay is not feasible. In the next chapter we address numerous innovative schemes to detect position directly between a mask and substrate with interferometric sensitivity.

1.5 Thesis Outline

The goal of this thesis is to invent a broadly applicable position metrology method that achieves sub-nanometer detectivity and nanometer overlay of a mask and substrate and operates irrespective of environmental disturbances. Chapter 2 provides background on a variety of next-generation alignment methods. Chapter 3 introduces the fundamentals of a novel positioning scheme, called Interferometric-Spatial-Phase Imaging (ISPI). Chapter 4 describes several novel means of detecting gap, including one with long-range interferometric sensitivity, and an oblique-incidence Fabry-Perot interferometer. Chapter 5 introduces a gap detection method with sub-nanometer sensitivity, called Transverse Chirp Gapping (TCG), based on frequency and phase encoding of gap in interference fringes. Chapter 6 introduces a new means of verifying gap using near-field self-imaging of the same TCG mark. Chapter 7 describes the effects of perturbations on ISPI, including process layers, variations in the viewing angle of an ISPI microscope, axial camera rotation, and grating defects. Chapter 8 describes the design and construction of the ISPI alignment apparatus and source alignment method and apparatus. Chapter 9 presents experimentally-demonstrated sub-nanometer detectivity and nanometer-level overlay using ISPI, including exposures in resist using feedback-stabilized alignment and gate-to-gate alignment in a double-gate transistor. Chapter 10 summarizes and concludes the dissertation.

2 Overview of Alignment Methods

In the past two and a half decades, the promise of higher lithographic resolution via shorter wavelengths spurred development of new generations of alignment methods at universities and industrial laboratories. Linewidths of 30 nm were fabricated using a 1.3 nm, soft x-ray wavelength [2.1, 2.2], about an order of magnitude increase beyond conventional lithographic resolution using ultraviolet light. To make devices that could take advantage of this increase in resolution, pattern overlay requirements decreased to <10 nm. Many alignment schemes were developed, with the common factor being an attempt to adapt the success of stage-control interferometry to direct, interferometric measurements between the mask and substrate.

The challenge of direct interferometric referencing of mask and substrate was made evident by the problems that arose when subjecting the substrate-related part of the interferometer to process-induced variations. Under ideal conditions on a test bench,

many of these alignment schemes could sense sub-10 nm mask-substrate displacements, but in typical manufacturing conditions, overlay was corrupted by process-induced pattern variations. Some of the most common sources of error were layers of resist, uneven resist flow over the marks, biasing induced by etching and deposition, variations in diffraction efficiency of the substrate marks, and changes in the index of refraction within the interferometer beampaths, due to density variations and helium/air mixing. Other problems included laser intensity fluctuations and very limited acquisition range. In overview, the high sensitivity implied by interferometry was frequently undermined by encoding alignment in the *amplitude* of the interferometric signals, which made them susceptible to noise and perturbations, analogous to reception corruption in AM radio.

2.1 Optical Alignment

For at least the first two decades of the semiconductor industry, optical alignment was used exclusively. The principles of optical resolution were well established by the late 1890's. Optical resolution is characterized by the minimum resolvable period in a grating[8], p_{min}

$$p_{min} = \frac{\lambda}{2NA} \tag{2.1}$$

where λ is the wavelength of illumination and NA is the numerical aperture. NA is defined as

$$NA = n\sin\theta \tag{2.2}$$

where n is the index of refraction of the medium between the microscope objective and the imaged object and θ is the maximum half-angle over which light can be collected by

[8] We consider any feature on a sample to be a Fourier composition of periodic components. Resolution is characterized by the highest spatial frequencies in the Fourier series.

the objective lens. For visible light, in the range between 400 and 700 nm, a high-quality optical microscope with an NA of 0.95 has minimum resolution $p_{min} \approx 200$ nm[9].

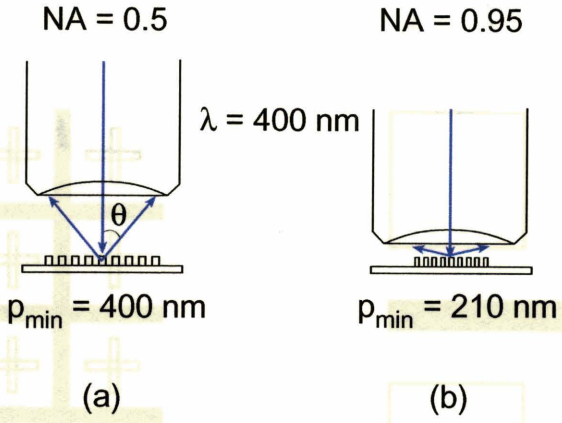


Fig. 2.1. Lens geometries for two values of NA. The working distance of the lens decreases with increasing NA to capture the beams diffracted at larger angles. According to the Abbé theory of imaging, at least two orders must be captured to reproduce the fundamental grating period in the image plane of the microscope.

2.1.1 Suss MicroTec

In industrial applications, a common optical aligner is the one manufactured by Suss MicroTec [2.3], which uses three high-NA (NA = 0.55), normal-incidence optical microscopes with white-light illumination. The high NA of the microscopes implied a short depth of focus (±1 μm), which could be problematic, since the mask-substrate gap was typically 50 μm. The mask and substrate had to be brought into focus separately, precluding simultaneous alignment measurement, and allowing drift, latency, and optical axis deviations to contribute to alignment errors.

The Suss microscopes detected alignment from marks that were an evolutionary step beyond the historically-common cross-and-box marks. In the cross and box marks

[9] Resolution will be improved by increasing the index of refraction, n, in the optical path between the objective and the sample. For instance, oil immersion (n = 1.4) reduces p_{min} to ~140 nm.

(Fig. 2.2(a)), the mask contained a cross, which was centered in X and Y between a set of four boxes in each of the four surrounding quadrants. The Suss alignment marks (Fig. 2.2(b)) consisted of orthogonal lines on the mask, with corresponding marks interspersed on the substrate.

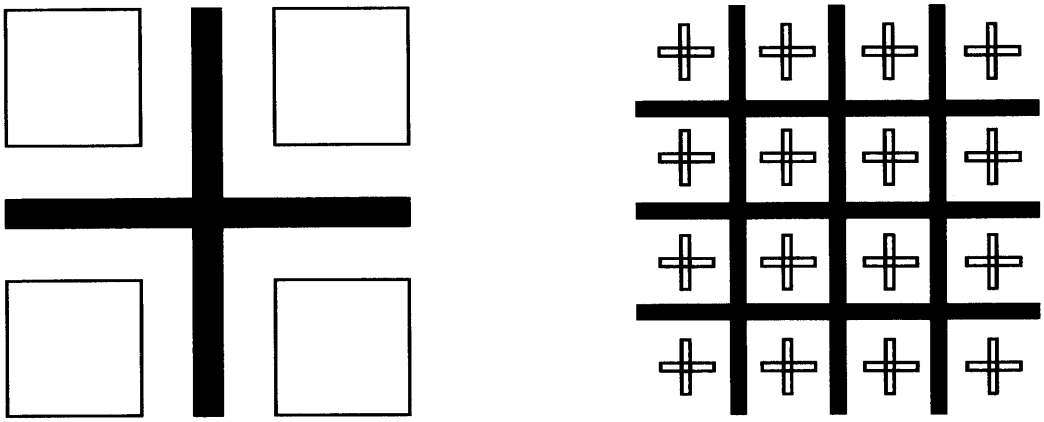


Fig. 2.2. (a) Cross and box alignment marks. (b) Karl Suss-type alignment marks. Dimensions of the mark were typically 200x200 μm.

CCD images of the mask and substrate alignment marks were acquired at two different focal positions and analyzed by a computer to determine alignment. In each image, the two-dimensional information was decomposed into one-dimensional signals by summing pixel values along rows or columns, producing a one-dimensional signal for X and another for Y. The one-dimensional signals contained sharp peaks corresponding to lines of the patterns. The relative position of the mask and substrate marks was determined with a peak-detection algorithm. Due to sub-pixel peak estimation and averaging over several peaks, alignment sensitivity exceeded the resolution of the microscope. Optical resolution of the Suss microscopes was ~360 nm, but published Suss alignment sensitivity was ~50 nm [2.4]. However, when common coatings such as resist were applied to the substrate, the demonstrated overlay was reduced considerably. Tests with five different types of materials (resist, polysilicon, and various metals) covering the substrate resulted in overlay errors with mean+3σ = 160 nm. Actual overlay of device features in mix-and-match optical and x-ray exposures of a MOS device

displayed overlay error of 300 nm [2.5]. During lithographic exposure the microscopes had to be removed from their observation position above the mask, allowing unobservable stage drift to contribute to alignment errors.

2.1.2 Miyatake and Hirose

Another optical technique, called the Scattered Light Aligner (SLA) [2.6, 2.7, 2.8, 2.9], was developed by Miyatake and Hirose at Sumitomo Heavy Industries. The motivation was to move the microscope objectives out of the way of an exposing beam and detect alignment during exposure. In the SLA scheme a microscope observed the minute fraction of light scattered at an oblique angle from the microscope-facing edges of rectangular marks on the mask and substrate. A schematic of the SLA scheme is shown in Fig. 2.3.

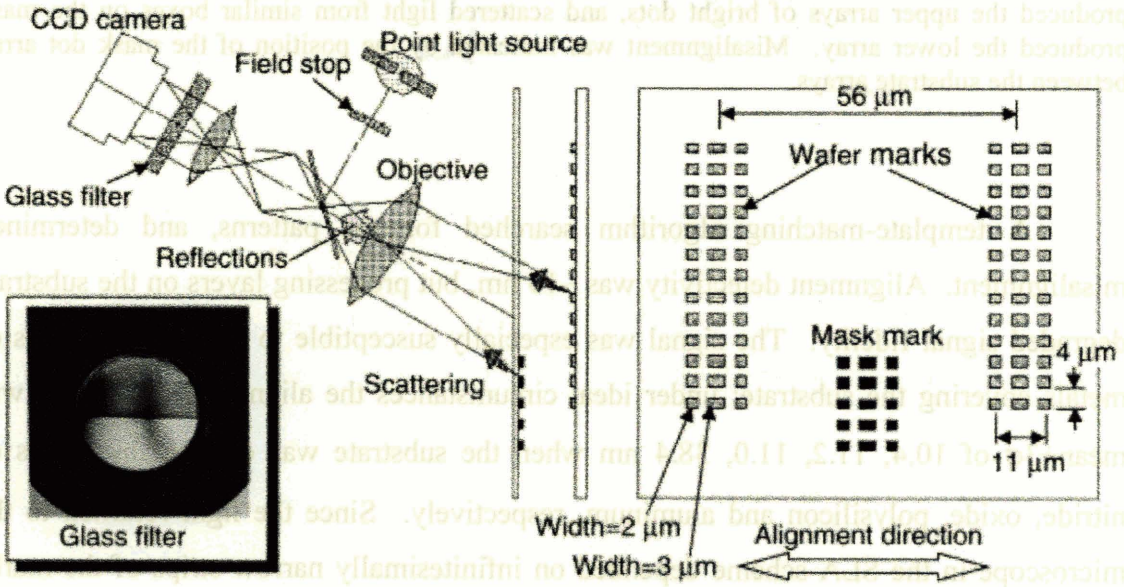


Fig. 2.3. Schematic of the Scattered Light Aligner. The alignment signal consisted of light scattered from arrays of boxes on the mask and substrate. The lower half of the glass filter had a neutral-density filter, which reduced the amount of light returning from the mask scattering, and equalized the intensity of light imaged from the mask and substrate.

Broadband light from a 100 W, super-high-pressure mercury arc lamp was delivered to an optical microscope via a multimode fiber-optic bundle, and scattered from the facets of an array of boxes facing the microscope. One set of boxes was on the mask and a surrounding pair of box arrays was on the substrate. A 20x microscope collected the scattered light at an angle of 30° from the mask normal, where it was detected by a CCD camera. A resulting image is shown in Fig. 2.4.

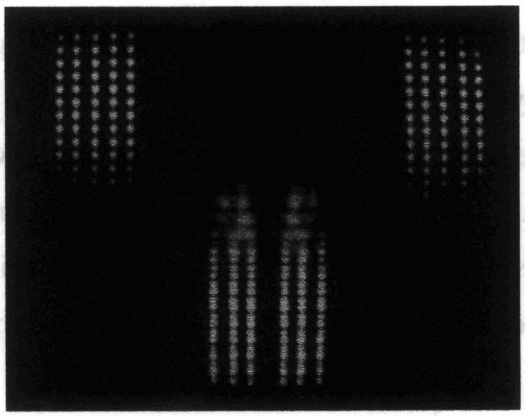


Fig. 2.4. Image from the SLA aligner. White light scattered from boxes on the substrate produced the upper arrays of bright dots, and scattered light from similar boxes on the mask produced the lower array. Misalignment was indicated by the position of the mask dot array between the substrate arrays.

A template-matching algorithm searched for the patterns, and determined misalignment. Alignment detectivity was ~10 nm, but processing layers on the substrate degraded signal fidelity. The signal was especially susceptible to surface roughness of metals covering the substrate: under ideal circumstances the alignment detectivity was mean+3σ of 10.4, 11.2, 11.0, 38.4 nm when the substrate was covered by layers of nitride, oxide, polysilicon and aluminum, respectively. Since the light returned to the microscope in the SLA scheme depended on infinitesimally narrow strips of the marks perpendicular to the incident illumination, it was very sensitive to processing and line edge roughness. Any process step that affected the line edge roughness, or changed the line-edge curvature, or slope, had the potential to affect the alignment measurement.

2.2 *Alignment via Linear Zone-Plate*

2.2.1 *Fay, Trotel, and Frichet*

In the Fay, Trotel, and Frichet scheme [2.10], a linear zone plate focused HeNe laser light onto a substrate, creating a narrow swath of light parallel to a grating on the substrate (Fig. 2.5).

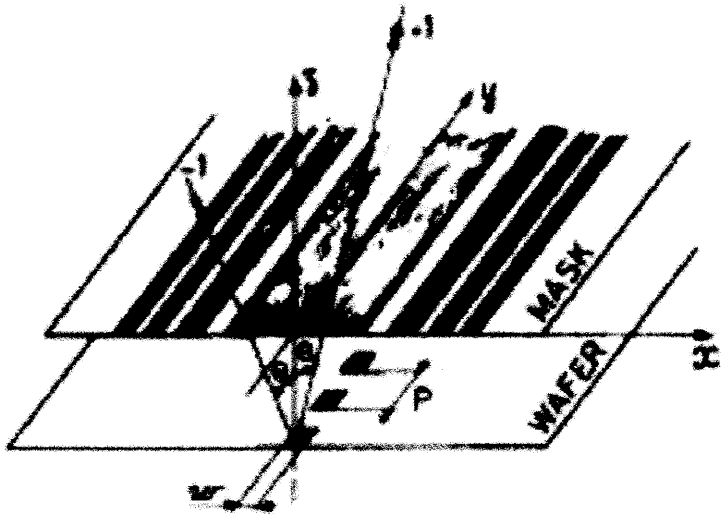


Fig. 2.5. Alignment scheme by Fay, et al. using a linear zone plate on a mask and a narrow, orthogonal diffraction grating on a substrate. A normal-incidence laser beam illuminated the marks while the zone plate was scanned over the substrate grating. Alignment was signified by a voltage peak detected at a photodiode. The intensity of the diffracted beam was sensitive to gap due to the small focal depth of the zone plate.

When the mask zone plate was scanned over the substrate grating, a portion of the focused beam diffracted in a plane orthogonal to the k-vectors of the zone plate[10]. The

[10] Alternatively, the angle of the incident beam could be rotated through small angles in the plane containing the normal and the k-vectors of the linear zone plate. The range of motion of the laser spot on the zone plate was ±1 μm. Scanning the mask yielded a sharper convolution peak, but scanning the incident angle was faster. The mean scan angle could be slightly oblique to the mask normal without aversely affecting the focusing property of the zone plate, thereby moving the optics out of the exposing beampath.

intensity of the diffracted beam, as measured with a photodetector, reached a peak when the center of the zone plate was aligned over the center of the substrate grating. The shape of the intensity signal as a function of distance was a convolution of the grating width and the focal spot width. Although theoretical sensitivity was ~0.5 nm, the best experimentally determined alignment sensitivity was ~50 nm.

The intensity of the detected beam was sensitive not only to alignment, but also to gap, with the intensity maximum occurring at a gap equal to the focal depth of the zone plate. Gap could be detected to ~1 μm by moving the mask through the focal plane of the zone plate.

Some of the primary disadvantages of this scheme were the coupling of alignment and gap in a single intensity, the time necessary to perform a scan, and the relatively poor alignment sensitivity achieved in practice. The need to scan meant that an intensity measurement acquired in a single moment was not indicative of alignment. Any scheme that requires scanning the mask to detect alignment (or gap) places demands on the stages to accurately move back to a calculated position, and over time the stage position will have a tendency to drift relative to the alignment position. This scheme was also lacking in generality since it required working at a single, predetermined gap to avoid diminution of intensity at gaps not corresponding to the focal depth of the zone plate.

2.2.2 Chen and Cerrina

To avoid difficulties in the Fay et al. scheme resulting from scanning the mask or the angle of the incident illumination, Chen and Cerrina devised an alignment method called a "two-state aligner" or TSA [2.11]. As in the Fay et al. scheme, light incident on a linear zone plate on the mask was focused on periodic features on the substrate. An electro-optic modulator switched the focused beam between two sides of the mark, and a

photodiode detected the intensity of light diffracted from the two sides of the mark. Alignment was indicated by the relative intensity of the diffracted light returning from each of the two positions, thereby indicating the relative position of the dots with respect to the mask zone plate. The depth of focus of the zone plate was ±2 μm. The detectivity of the system was estimated to be 2 nm, but it was found to be susceptible to errors from a layer of polysilicon covering the substrate marks. The substrate coatings distorted the measurements with a 3σ = 60 nm [2.12]. In experiments, the overlay of device features was 3σ = 180 nm [2.13]. The scheme also suffered from a small range: alignment could be detected over a range of only ±200 nm (along the X direction).

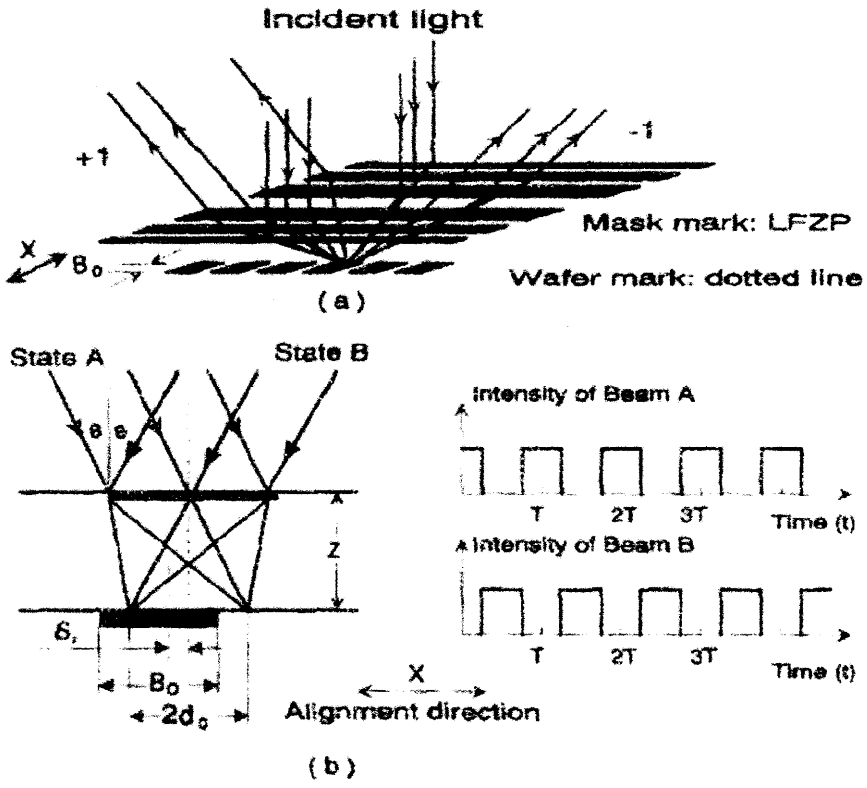


Fig. 2.6. Schematic of the two-state aligner.

2.2.3 Zhou and Feldman

To avoid the need to electro-optically alternate positions of the incident beam on the zone plate, Zhou and Feldman used signal differences between a pair of gratings on the substrate [2.14]. In one configuration they used two slightly different periods, and in another configuration they used the same grating period with a relative π-phase shift, as shown in Fig. 2.7. The gratings were illuminated with a 1 μm-wide stripe of light from a zone plate on the mask. The angle of illumination was scanned over a small angle with a rotating mirror to obtain a time dependent signal. The scanned light swept over the zone plate along the zone plate's k-vector direction, causing the focused beam to sweep over the substrate gratings, perpendicular to the interface between the two gratings. The first order diffractions from each of the gratings were detected by a split photodiode. The difference in voltages from each half of the photodiode indicated alignment. In the case of the π-phase-shifted gratings, alignment was indicated by an intensity minimum during the laser scan. In the case of the dual grating periods, alignment was found from the point where the intensity plots of both halves of the split photodiode had a maximum slope.

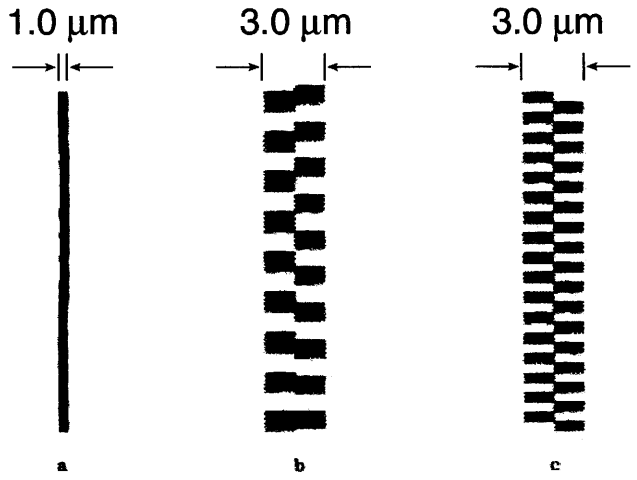


Fig. 2.7. Gratings used by Zhou and Feldman. A linear zone plate focused light into a 1.0 mm wide stripe (a). Light was diffracted at slightly different angles by two gratings of period 3.8 and 4.0 mm (b) or by two gratings of the same 2.0 mm period with a p-phase shift between the gratings(c). The width of each grating pair was 3.0 mm, which determined the capture range.

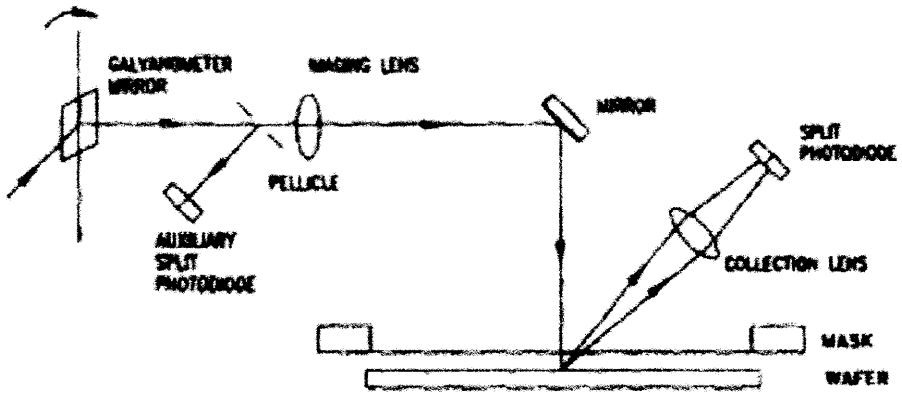


Fig. 2.8. Apparatus of Zhou and Feldman. The linear zone plate was illuminated by a HeNe laser. A galvanometer mirror scanned the incident beam through a range of angles at near-normal incidence. Light diffracted by gratings on the substrate was imaged onto a split photodiode. The difference signal in the split photodiode indicated alignment. Instability in the emission direction of the laser amounted to 10 μrad, and necessitated detection of the incident beam direction with an auxiliary photodetector.

The sign of the slope of the intensity plot in either variant of this scheme indicated the direction necessary to move toward the aligned position, but it did not encode the degree of misalignment, i.e., the displacement required to achieve alignment. Instead, the substrate must be scanned to find the aligned position, and then moved to the stage position correlated with the alignment condition. In other words, it is a non-metrological system, since it finds a point of symmetry, but does not measure displacement. In many situations, a best-fit alignment position requires deliberate misalignment at several alignment marks in order to achieve an optimal fit between the mask and substrate.

Measurement sensitivity was exceptional; without resist or other process coatings 3σ = 1.8 nm for the dual-period gratings, and 3σ = 0.7 nm for the π-phase-shifted gratings. However, overlay of exposed features was not reported, and the beam steering mirror (Fig. 2.8) for the incident laser illumination blocked the exposing beam, precluding feedback control throughout exposure.

2.3 *Interferometric Intensity Alignment*

2.3.1 *Flanders, Smith, and Austin*

Flanders, Smith and Austin pioneered interference of diffracted beams for mask-substrate alignment [2.15]. In this scheme, alignment was indicated by the relative intensity of symmetric, diffracted beams. Superposed gratings on a mask (mask period p_m = 1 μm) and substrate (substrate period p_s = 1 μm) were illuminated at normal incidence with a HeNe laser, as illustrated in Fig. 2.9.

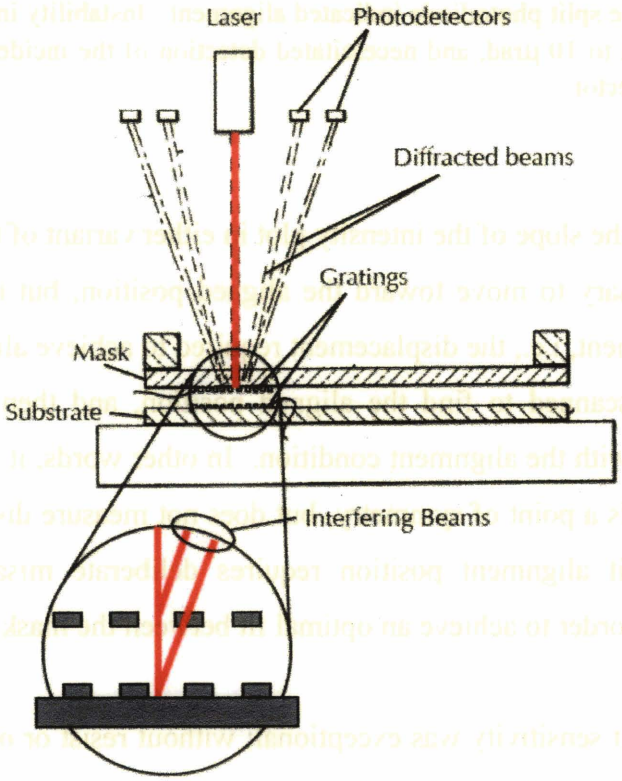


Fig. 2.9. Alignment scheme of Flanders, Smith, and Austin. Similar backdiffracted orders (e.g., +1 from the mask and +1 from the substrate) interfere. The interferometric intensity was measured by photodetectors. Alignment was indicated by the voltage difference between two photodetectors measuring symmetric (+1 and -1) diffracted order groups.

Each grating produced backdiffracted orders which interfered along common paths above the mask plane. The phase difference $\Delta\phi_m$ between corresponding orders was given by

$$\Delta\phi_m = \frac{2\pi m}{p}\Delta x \tag{2.3}$$

where m was the diffracted order, p was the grating period, and Δx was the displacement.

The intensity of the interference within each set of similar orders was transduced to a voltage using a photodetector above the mask. The arrangement was symmetric for +1 and -1 orders, and the difference between the intensities of +1 and -1 order groups indicated alignment. A zero voltage difference between the photodetectors would ideally indicate the aligned position. However, the intensity at the photodetectors was affected by factors such as blazing of the substrate mark during a sequence of process steps [2.16] and variation in the response of individual photodetectors. Such error sources made it difficult to obtain a linear relation between detected voltage and misalignment, and could bias the point of zero voltage difference. In practice, 20 nm overlay was achieved[11].

The intensity was also dependent on gap, and tilt of the mask. Coupling between gap and alignment complicated measurements, and necessitated fine leveling prior to alignment. The mask was leveled within 1 μm by observing interference fringes in the vicinity of the photodetectors due to an angle between like-order beams diffracted from the mask and substrate. Rotation between the mask and substrate was determined from a similar interference effect. The alignment signal dropped to zero at zero gap, and at periodic finite gaps where the phase difference between beams within a diffracted group went to zero. Measurement range was determined by the period of the gratings, where identical intensity was found upon every displacement by ±1/4 the grating period. Coarse alignment required an auxiliary method.

[11] In a later version, used to achieve this result, the multiple photodiodes were replaced with a single photodiode and a set of mirrors, with a chopper to time-multiplex the various interferometric signals on the single photodiode.

2.3.2 Nelson, Kreuzer, and Gallatin

Silicon Valley Group Lithography (SVGL) developed a through-the-mask (TTM) alignment scheme [2.17] based on alignment amplitude measurements made during a scan of the substrate stage under the mask. In the TTM scheme, a laser beam containing several wavelengths between 780 and 870 nm was incident at a 6° angle on a set of marks on the mask and substrate, consisting of a linear grating and a checkerboard, respectively (Fig. 2.10). The interference of a set of four beams, diffracted back at 18°, was detected by a photodiode. The intensity on the photodiode oscillated sinusoidally within an envelope determined by the overlap between the mask and substrate marks during the scan. The aligned position was interpreted to be the maximum intensity of the envelope.

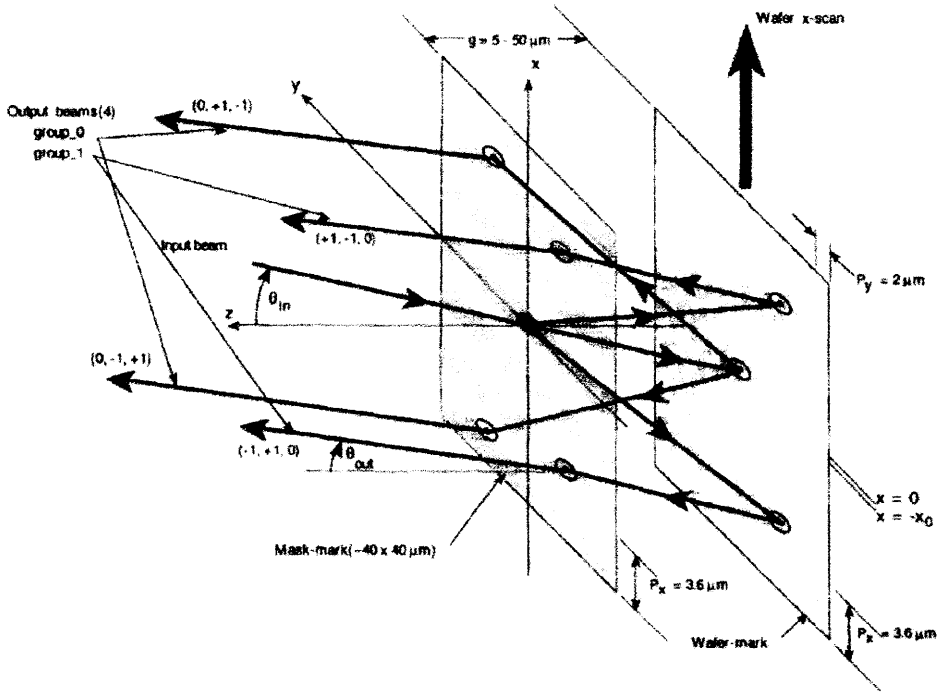


Fig. 2.10. Mark diffraction paths from TTM design.

During a scan of the substrate, the maximum intensity was correlated with the position of the substrate stage. After the scan, the stage moved back to the calculated aligned position, based on position measurements from the stage interferometer.

Alignment accuracy with TTM was found to be $3\sigma = 40$ nm. An 8 nm component of the error was attributed to resist-induced variations.

An advantage of the TTM scheme was that it could detect alignment without blocking the exposure path, but a major problem was that alignment *could not be detected without scanning* the mask or substrate, i.e., a movement was required to detect a position. Movement after measurement placed additional burdens on the stages - any stage error in the move-back position appeared as an overlay error. Since the stage had to be scanned to detect alignment, there was no way to verify alignment after the movement to the calculated position, short of scanning the stage again, which would merely reiterate the uncertainties in stage position. Even if the stage error was initially zero, thermal expansion and drift within the mask and substrate stages would not necessarily be detected with adequate resolution by the laser interferometer at the edge of the stage, and could cause deviations from the desired mask-substrate position.

Other aspects of this scheme were that X and Y measurement required separate scans of the substrate, and gap was not detectable.

2.4 Heterodyne Alignment

2.4.1 Suzuki and Une

Suzuki and Une at NTT LSI Laboratories developed an aligner based on heterodyne of two laser signals [2.18]. The heterodyne produced a phase signal from interference of two laser frequencies, which was detected with a photodetector. Two orthogonally polarized beams of frequencies f_1 and f_2 were incident on a constant-period grating. The beams were incident at two different oblique angles to the grating, arranged so that the +1 and -1 diffracted beams were diffracted at the same angle (parallel to the

surface normal), and interfered, as illustrate schematically in Fig. 2.11. Noise-limited sensitivity of the alignment signal was ~16 nm.

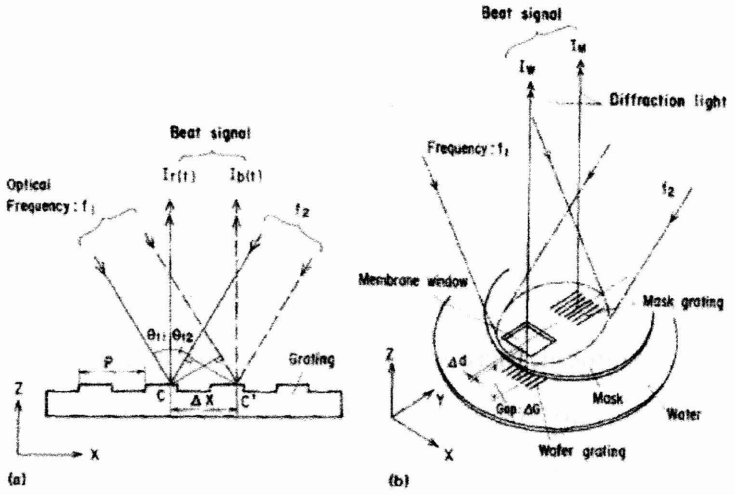


Fig. 2.11. Alignment scheme of Suzuki and Une. Alignment was detected from the beat signal between two diffracted first order beams from incident laser beams of frequencies f_1 and f_2. To detect alignment a second grating of the same period on the substrate was illuminated through a window in the mask. Two beat signals were detected along the Z axis with two photodetectors.

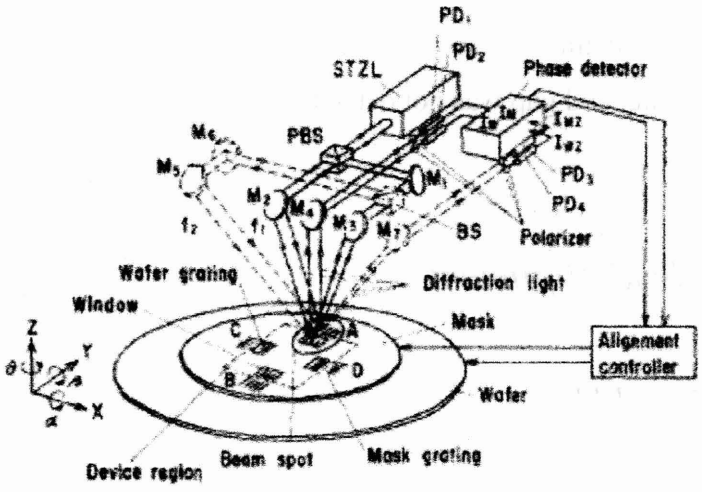


Fig. 2.12. Alignment apparatus of Suzuki and Une. The mask had a grating of 4 mm period. Marks were 200 mm long. The laser source was a frequency-stabilized HeNe with Zeeman line splitting. Mirrors M_2 and M_3 directed the f_1 and f_2 beams onto the mask and substrate marks. The beat frequencies of the recombined beams were reflected by mirror M_4 and detected by two photodetectors PD_1 and PD_2. The mask and substrate gratings were separated sufficiently to isolate one heterodyne signal from the mask and substrate onto each photodetector. Gap was detected with interference of asymmetric higher-order diffracted beams, along paths from mirrors M_7 and M_8.

One difficulty in this scheme was that process coatings affected only one of the two heterodyne signals, thereby causing alignment deviations. The best overlay accuracy obtained under manufacturing conditions was $3\sigma = 130$ nm.

In this alignment scheme, an auxiliary scheme, such as an optical pre-aligner, was needed to overcome the phase ambiguity that occurred with every intensity cycle. The problem of phase ambiguity was compounded in gap signals, which were mixed with an alignment-dependent phase.

2.4.2 Itoh and Kanayama

In the alignment scheme proposed by Itoh, et al. [2.19], two gratings were on the mask, each having the same period $p = 1.0$ μm.

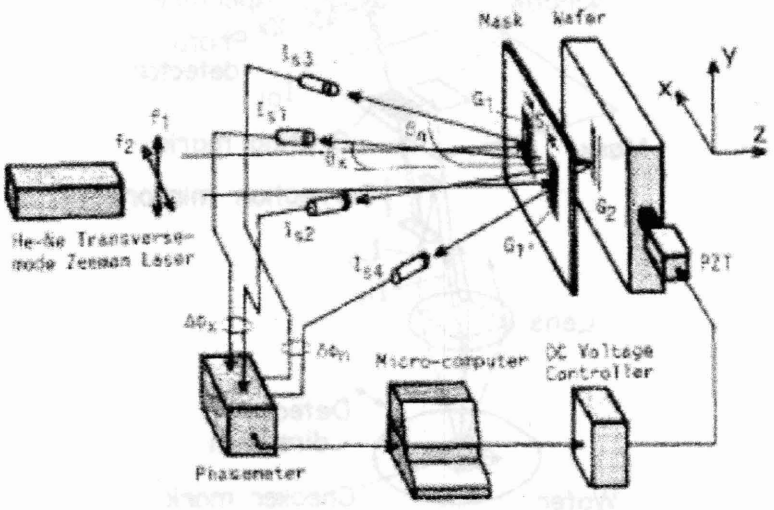


Fig. 2.13. Apparatus of Itoh and Kanayama.

A single grating was on the substrate, having period $1.5xp = 1.5$ μm. The gratings were illuminated at normal incidence with a two-frequency HeNe laser, producing four

interference beams, detected by four photodiodes, each with a beat frequency of $|f_1-f_2|$. The alignment signal was detected by the relative intensity of the innermost sensors, I_{s1} and I_{s2}, and was claimed to be "minimally affected" by variations in gap. Mask tilt was critical to achieving accurate alignment. Tilt was detected by sensors I_{s3} and I_{s4}, and corrected prior to alignment. Alignment was detected within 20 nm.

2.4.3 Tabata and Tojo

Tabata and Tojo [2.20] described an alignment system based on diffraction of light from a checkerboard on the substrate and a linear grating on the mask, illuminated by a HeNe laser, as shown in Fig. 2.14.

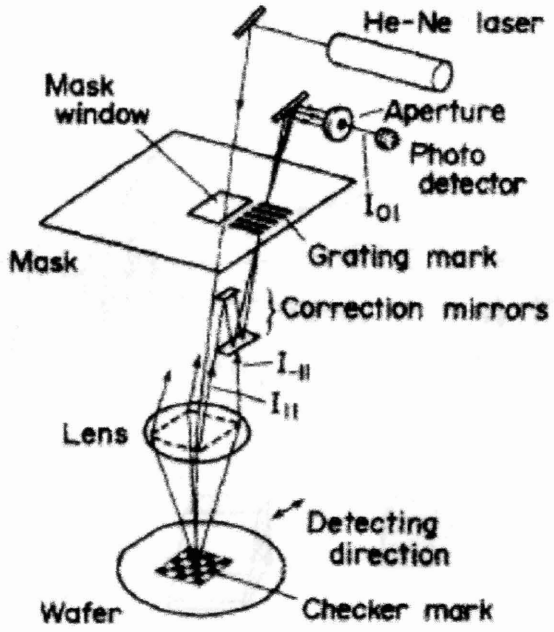


Fig. 2.14. Alignment scheme of Tabata and Tojo. A laser beam diffracted from a checkerboard on the substrate, and the interference of these first-order diffractions with a linear grating on the mask produced a spatial heterodyne signal. The zero-order diffraction from the mask grating was detected by a photodetector.

Laser light was incident directly on the checkerboard, through a clear window in the mask surface. Two 1st order diffracted beams from the checkerboard passed through a lens, and a linear grating on a mask modulated the interference of two such diffracted beams, producing a spatial heterodyne signal[12]. A photodetector picked up a portion of the interfering light, resulting in an intensity signal that varied sinusoidally as a function of the mask displacement along the k-vector of the mask grating.

The range of measurements was restricted to a ±1/4 of the grating period, and the minimum detectivity was

$$\Delta x = \frac{\phi}{2\pi} p \tag{2.4}$$

where ϕ was the minimum detectable phase difference and p was the grating period. Position estimation was limited by noise in the phase detection.

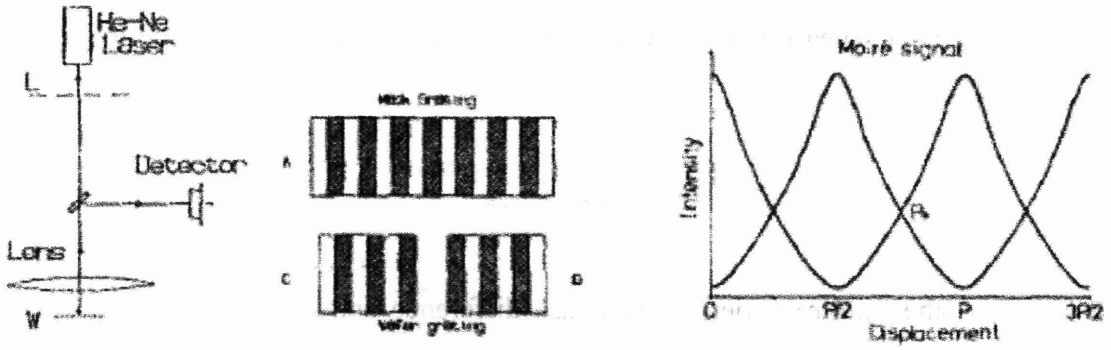
Since the intensity of the alignment signal was periodic with displacement, a quadrature scheme was used to determine the direction of displacement. In this case a second checkerboard and grating were used adjacent to the first grating, the difference being the second grating was shifted by a quarter period with respect to the first, providing two sinusoidal signals in quadrature. The incident beam was switched between marks using an acousto-optic deflector.

Optical signal reduction caused by processing layers was a problem, but an auto-gain circuit was used to compensate. With the auto-gain circuit, the errors in alignment caused by processing layers were reduced to ±40 nm. In practice, overlay of 150 nm was achieved.

[12] The spatial heterodyne could also be considered a moiré – however, in this scheme, the photodetector samples the intensity at only a single point, unlike image-based moiré schemes.

2.4.4 Hara and Uchida

Using a spatial heterodyne, two gratings of period p_1 and p_2 were illuminated at normal incidence with a HeNe laser [2.21]. A reflected component was detected by a photodetector.



(a) (b) (c)

Fig. 2.15. Experimental setup (a) and grating patterns (b) used by Hara and Uchida. Gratings with 25 μm period were on the mask and substrate. A 180° phase shift was introduced between two parts of the substrate grating. The intensity of a signal was detected during a scan of the substrate.

Reflection from the substrate caused undesirable interference effects. In addition, the mask and substrate formed a Fabry-Perot cavity so the signal intensity depended strongly on the gap. To avoid inter-reflections, the mask was tilted by 0.1° with respect to the substrate plane.

A constant-period grating was on the mask and a different grating period was used on the substrate mark. The substrate grating was divided into two parts, with a π-phase shift between each part. The result was an intensity signal detected by two photodiodes, one above each substrate grating. The difference in the intensities indicated the sum of two π-phase-shifted intensity oscillations. The aligned condition was at the

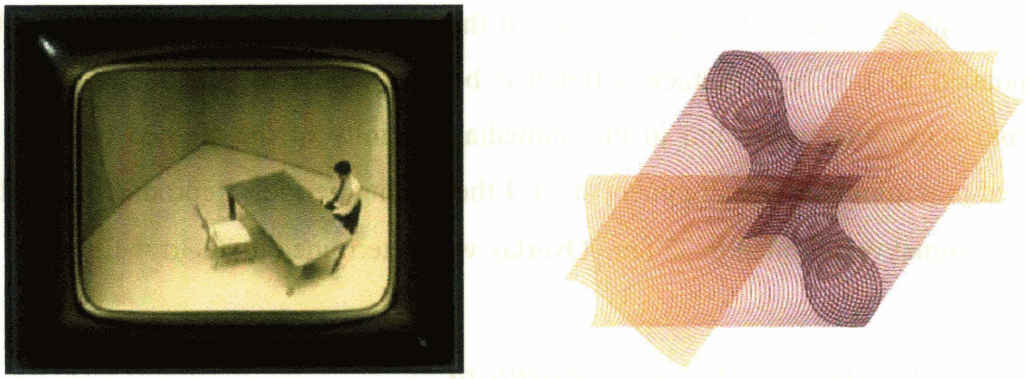
region of highest slope in the superposition of the two oscillations (Fig. 2.15(c)), which corresponded to the zero voltage difference between the two photodetectors. The maximum sensitivity was 5 nm in the immediate vicinity of the aligned position and decreased to a minimum misalignment at ±1/4 the grating period. Reproducibility of the alignment signal was limited to 32 nm. Overlay was determined to be $3\sigma = 120$ nm.

2.5 Interferometric Moiré Alignment

Moiré aligners retain the advantages of two-dimensional imaging of optical aligners in Section 2.1 (i.e., X, Y, and θ measured at the same time), do not require scanning (a requirement of the zone-plate schemes in Section 2.2), and can match or exceed the sensitivity demonstrated by the interferometric and heterodyne aligners in Sections 2.3 and 2.4.

2.5.1 The Moiré Phenomenon

The ethereal stripes of the moiré phenomenon may be mundane, yet never fail to attract and fascinate, due to their restless complexity of shape and extraordinarily rapid motion. Some form of moiré fringes are visible almost anywhere, whether in overlapping folds of a handkerchief or a shower curtain, in a screen door, between guardrails on a bridge, in striped clothing on television, or in digital photographs of television screens (Fig. 2.16(a)). In these cases the moiré patterns are accidental and convey no useful information. Other moiré patterns are created intentionally, but are generally used for their aesthetic attributes, such as in art (Fig. 2.16(b)), novelty items (a pair of perforated cards moved past each other to create an illusion of motion, e.g., a butterfly flapping its wings), or clothing. The term moiré originated with "watered" silk, known for its shimmering, wavy fringe patterns, which respond to the wearer's movements.



(a) (b)

Fig. 2.16. Examples of moiré: (a) Sampling moiré, such as obtained by taking a digital photograph of a TV screen. (b) Moiré formed by overlapping S-curves.

Although moiré fringes have been a common occurrence as long as periodic structures were in existence, it was not until the late nineteenth century that Lord Raleigh found a scientific application for moiré. Raleigh's pursuit of spectral analysis of light led him to an interest in ruling diffraction gratings. At the time, the grooves of diffraction gratings were inscribed in a substrate with a sharp tip, under the control of a mechanical ruling engine. However, imperfections in the gratings were the inevitable result of nonlinearities in drive mechanisms and stage motion, causing deviations from parallelism and deviations in spacing between grating lines. Periodic errors in the grating period, tied to the leadscrew rotation period, were most common. Such imperfections resulted in spurious diffracted orders that corrupted spectroscopic measurements[13]. Searching for a means to evaluate grating quality, Raleigh found in 1874 that the superposition of two gratings produced a set of low-frequency fringes that varied in accordance with the relative quality of the gratings [2.22], [2.23]. Raleigh noted that if the two gratings were of the same period, the fringes would be visible only if one grating was rotated; alternatively, existence of the fringes when the gratings were not rotated indicated disparity in the grating periods. Here was a simple, intuitive, yet sensitive means of comparing the quality of ruled gratings.

[13] Random errors caused broadening of spectral lines, but of more concern were periodic errors that resulted in additional, spurious spectral lines known as Rowland's ghosts.

Despite Raleigh's recognition of the usefulness of moiré patterns, wider scientific and engineering uses were not developed for many years. The concept of measuring displacement with moiré patterns was not proposed until 1887 (Righi), and the first application was delayed until 1927 (Giambiasi). Despite the sensitivity of moiré, very little use was made of moiré in displacement measurement until the 1950's when Guild [2.24, 2.25] applied moiré to machine control and devised optical encoders based on moiré to measure linear translation of stages. The principle of moiré encoding has widespread use in contemporary linear encoders, although grating quality presently limits their accuracy to a few tens of nanometers[14].

Moiré patterns are of significance to this dissertation because they are *highly* sensitive indicators of the relative position of underlying structures. In addition, by virtue of their two-dimensional nature, moiré fringes convey far more information than can be detected by the intensity at a single point. For instance, multi-axis position information can be encoded in moiré fringes, and recovered, unlike the Fay et al. scheme, which combined both alignment and gap into a single intensity measurement.

2.5.2 King and Berry

The first use of moiré for mask-substrate alignment was by King and Berry [2.28]. In the King and Berry apparatus, moiré fringes were produced between circular gratings with periods of 4.0 and 3.8 μm, with a fringe period of 76.0 μm, including about 8 fringes in the 600 μm mark radius. The moiré fringes were observed with a 20x normal-incidence, bright-field, white-light microscope, with a 10x eyepiece. The gap between the mask and substrate was ~25 μm. Given that one fringe period was 76 μm,

[14] A recent advance in "ruling engines", the novel Scanning Beam Interference Lithography (SBIL), or "NanoRuler", writes a grating with an interferometric "brush", achieving high-speed writing and nanometer placement accuracy over a 300 mm diameter substrate [2.26, 2.27]. One application of gratings produced by the NanoRuler is to highly accurate linear encoders.

one would expect alignment accuracy by eye to be limited to about 1/20 of a fringe cycle, or about 4μm/20 = 200 nm. This is consistent with requirements stated by King and Berry for alignment of 1 μm pattern features, and matches their demonstrated alignment of 200 nm. A schematic of circular moiré fringes is shown in Fig. 2.17.

According to King and Berry the alignment of the marks is quantified in the following way: "For each cusp-like curve that passes through the common point, the mask and substrate are misaligned by one-half the pitch of the gratings. For example, if there are eight curves passing through the point, the pitch of the gratings is 4 microns, so alignment is off by 16 microns. To achieve alignment, we move the [substrate] along the axis of symmetry of the moiré pattern."

According to studies by the U.S. Air Force, the human eye is most sensitive to misalignment in circular patterns; hence circular moiré patterns are ideal from the perspective of a human operator. However, any industrial moiré alignment system requires automation using machine vision systems. Circular moiré patterns are not well matched to the rectilinear arrays of CCD cameras nor to e-beam grids used for patterning the alignment marks. Algorithms for analyzing circular fringes are also unnecessarily complicated, compared to those used to analyze linear fringes.

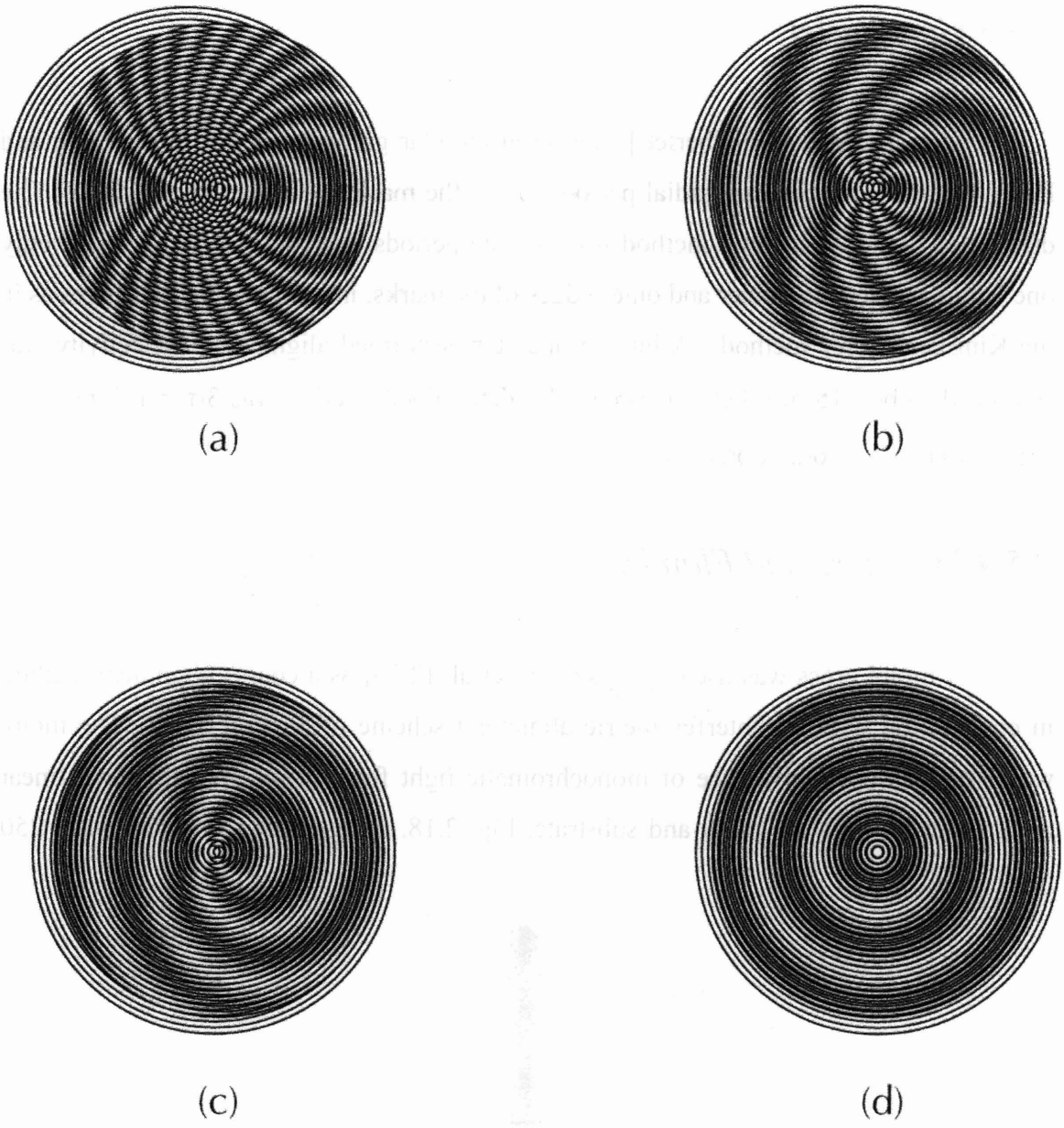


Fig. 2.17. Schematic of circular moiré patterns illustrating the method used by King and Berry, as well as Goodberlet. (a) Displacement in X by 3 grating periods, (b) Displacement in X by 1 grating period, (c) Displacement in X by 0.5 grating period, (d) Zero displacement. The period of one grating is 10% larger than that of the other grating. Circular moiré fringes effectively indicate the direction of offset (along the vector that symmetrically bisects the fringe pattern), as well as the magnitude of displacement. Displacement is proportional to the number of fringes that intersect the vector between the centers of each grating.

2.5.3 Goodberlet

More recently, Goodberlet [2.29] used circular gratings similar to the King and Berry method with constant radial periods, p_1 on the mask and p_2 on the substrate. The difference in the Goodberlet method was that the periods were designed to produce only one fringe between the inner and outer edges of the marks, instead of about 8 fringes as in the King and Berry method. A human operator performed alignment. Detectivity was estimated to be ~15 nm, but experimentally determined overlay was 3σ = 177 nm, due mainly to mask deformation in contact.

2.5.4 Lyszczarz and Flanders

A moiré cross was used by Lyszczarz et al. [2.30] as a coarse alignment method in conjunction with the interferometric alignment scheme of Section 2.3.1. The moiré was produced by interference of monochromatic light from identical, orthogonal linear zone plates on both the mask and substrate, Fig. 2.18. Alignment sensitivity was <250 nm.

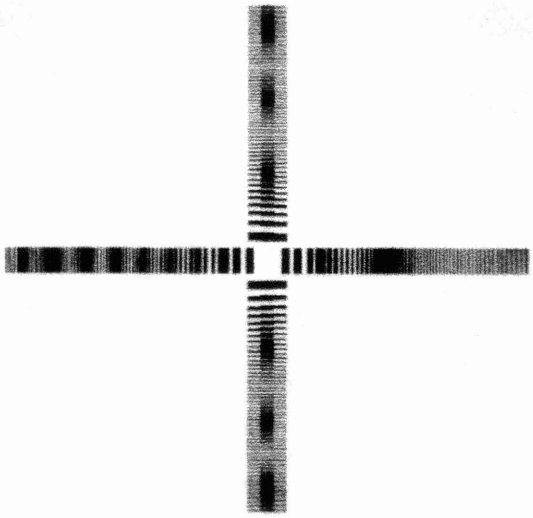


Fig. 2.18. Schematic of a moiré cross produced by a linear zone plate arrangement. Linear zone plates were identical on mask and substrate. The illustration shows alignment in the Y direction and misalignment in the X direction.

2.5.5 *Moel, Moon, and Smith*

Moel et al. used rectilinear moiré fringes compared to fiducial marks to facilitate automated analysis and e-beam generation of the moiré alignment marks [2.31]. In this scheme, called "On-Axis Interferometric" alignment, or OAI, a grating of period p_1 on a mask was superposed upon a grating of similar period p_2 on a substrate. A CCD camera with a 20x Leitz brightfield objective imaged interference fringes under white-light illumination, spatially collimated with a small-diameter (100 μm) pinhole in the back focal plane of the microscope.

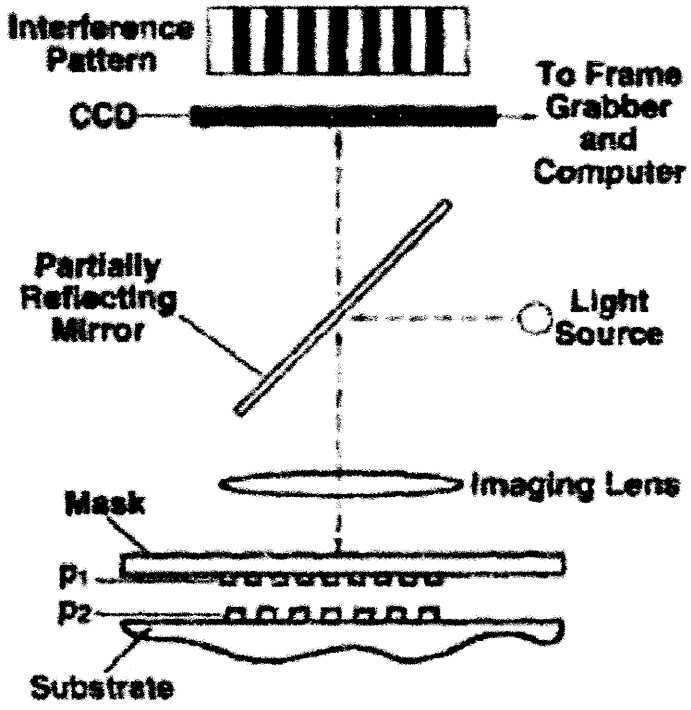


Fig. 2.19. Schematic of OAI alignment apparatus.

To determine alignment, the spatial phase difference was found between a set of moiré fringes, resulting from overlay of p_1 and p_2 gratings, and a coarse fiducial mark etched into the substrate, which was designed to have the same spatial period as the

fringes[15], as shown in Fig. 2.20. Additional sensitivity was obtained by analyzing the moiré fringes in the frequency domain. Alignment detectivity was < 10 nm.

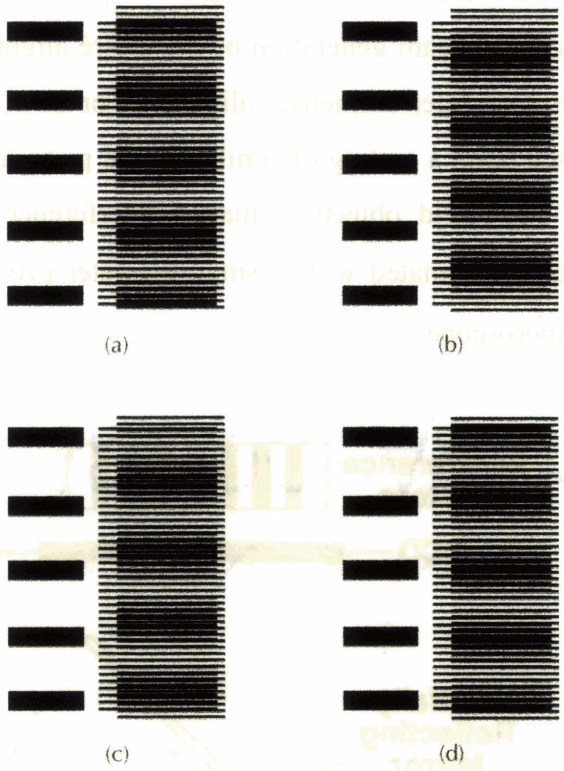


Fig. 2.20. Illustration of alignment marks used in the OAI alignment scheme. The substrate contains a set of fiducial marks and a grating. The mask contains only a grating. Displacements along the k-vector of the gratings are indicated in the sequence (a) – (d), corresponding to a shift of one period of the mask grating. Alignment was signified by the relative position of the set of moiré fringes, compared to a set of fiducial marks on the substrate. (a) Mask grating displaced by one grating period in +Y direction. (b) Displacement of half a grating period. (c) Displacement of a quarter grating period. (d) Aligned position. Note that displacement by one period realigns the moiré fringes to the fiducial marks. This ambiguity was resolved with conventional cross and box marks. Process layers, such as resist, affected the fiducial marks differently than the moiré fringes, inducing alignment errors under common conditions.

Although OAI demonstrated 3σ = 18 nm alignment in exposed features, it suffered from several disadvantages:

[15] The idea of using spatial frequency analysis in the analysis of optical systems was first attributed to Lohmann in 1957.

- The OAI microscope was normal to the mask, precluding the possibility of observing alignment at the critical mask-substrate position during exposure.
- In OAI, as in most practical systems, when alignment is not measured during exposure, alignment will drift before the exposure starts, causing unintended misalignment in the exposed features.
- Fringe contrast was very low due to specular reflections from the mask and substrate. In an attempt to compensate, gain was maximized in the camera, but such high gain increased electronic noise.
- Light from adjacent marks returned to the microscope and degraded the quality of the moiré fringes.
- The high-NA of the microscope resolved the grating lines, further degrading the moiré fringe visibility.
- The OAI microscope had a small depth of focus.
- The short working distance of the OAI microscope endangered the mask. A slight mis-focusing of the microscope would crash the objective lens into the mask, and destroy or damage the mask.
- Fringe contrast in the OAI scheme varied periodically with gap. Since the fringes were very faint to begin with, in order to have enough contrast to discern alignment, in practice it was found that alignment had to be performed at one of a discrete number of gaps. This restriction was undesirable, and reduced the generality of the scheme.
- The fiducial grating on the substrate was subject to alignment distortions caused by overlayers. The overlayers, such as photoresist or aluminum, had distinctly different effects on the fiducial pattern and the moiré fringes, resulting in misleading alignment measurements under most conditions encountered in lithography.
- In-plane rotation was measured by multiple iterations of the same microscope between two widely separated points on the mask, which is impractical in any commercial system.

- OAI detected X and Y alignment in the field of view of the microscope, but a general-purpose alignment scheme also needs to detect in-plane rotation, θ_z, and gap, Z, between the mask and substrate, as well as the out-of-plane rotations, θ_x and θ_y.

2.6 Summary

This chapter provided an overview of alignment schemes that progress beyond the simple box-and-cross optical scheme used since the beginnings of the semiconductor industry, and date back to the invention of the microscope. Interferometric techniques were applied with some success to this problem, but it proved to be a significant challenge to retain highly robust detectivity in the presence of a multitude of disturbances encountered on the substrate surface during a series of lithographic steps. Moiré schemes showed the most promise in this regard, but were limited by low fringe contrast and poor signal-to-noise ratio.

3 Interferometric-Spatial-Phase Imaging (ISPI)

This chapter introduces a novel alignment method called Interferometric Spatial Phase Imaging (ISPI). The central purpose of this method is to align a mask and substrate with sub-nanometer detectivity, yet in such a manner that the alignment information is unaffected by extraneous influences, such as the index of refraction of the medium in the beampath and process coatings covering the substrate alignment marks. To meet these goals ISPI encodes position in the spatial phase disparity between a matched pair of interferometric moiré patterns, viewed with low-NA, oblique-incidence, diffractive dark-field microscopes. Oblique-angle viewing eliminates the zero-order contribution to image formation, thereby providing high-contrast moiré fringes, and doubling sensitivity. In addition, oblique-angle viewing allows alignment to be measured and corrected at any time, before, and throughout the lithographic exposure. This chapter

will describe the specialized alignment marks, microscopes, illumination, fringe contrast, noise rejection, dynamic range, sensitivity, and feedback stabilization inherent to ISPI.

3.1 Fundamentals of ISPI

The ISPI method includes a) obliquely illuminating grating and checkerboard-type marks on a mask and substrate with collimated broad-band light, b) encoding relative mask-substrate position in fringes resulting from interference of multiple diffractions of the incident beam from the marks, c) acquiring images of the interference fringes with low-NA microscopes at a similar oblique angle, and d) decoding the position information from the relative phase of the fringes, using a spatial phase analysis algorithm. A schematic of a typical arrangement is illustrated in Fig. 3.1, and images from an ISPI microscope are shown in Fig. 3.2.

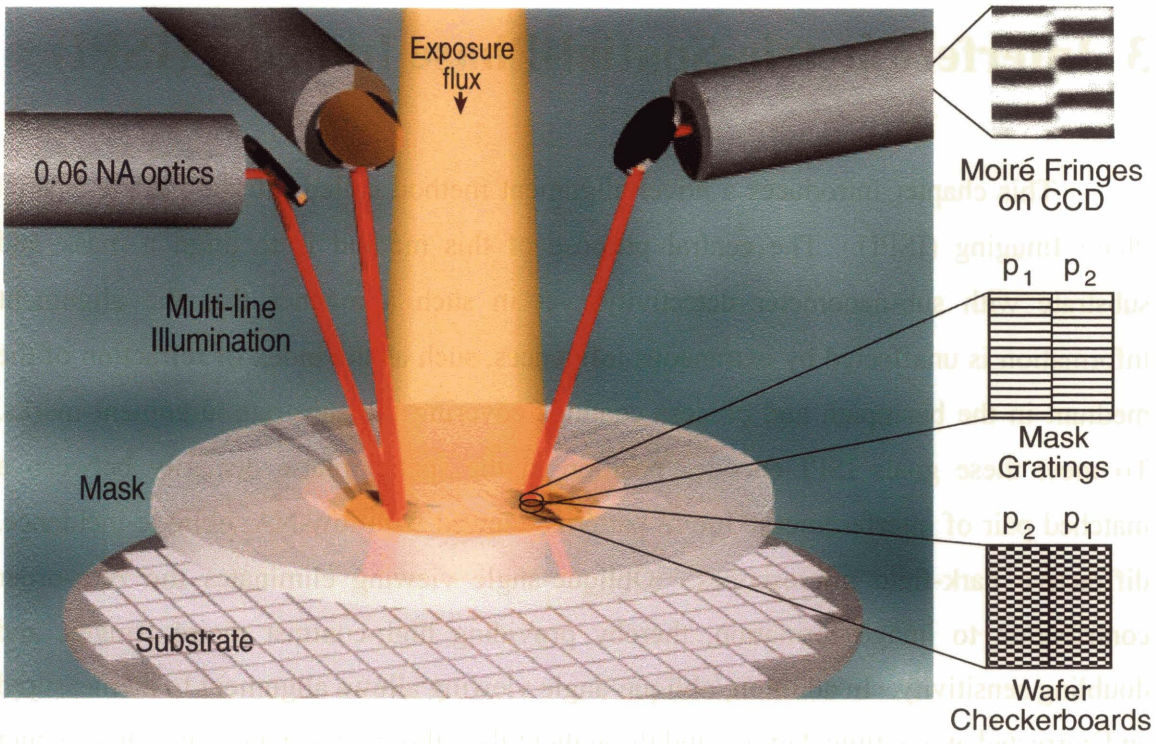
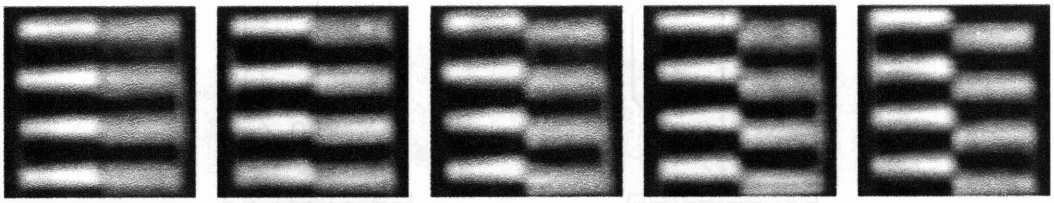


Fig. 3.1. Microscopes observe interference fringes produced by multiple diffractions from grating and checkerboard-type marks on two planar surfaces.



(a) (b) (c) (d) (e)

Fig. 3.2. ISPI alignment information is contained in the spatial-phase discontinuity of a pair of interference patterns. Images (a)-(e) illustrate spatial-phase disparities corresponding to steps of 50 nm.

3.2 Interferometric Moiré at an Oblique Angle

3.2.1 The Littrow Angle

In a conventional brightfield microscope, much of the light that returns to form the image comes from specular reflections, as shown in Fig. 3.3(a). In a conventional darkfield objective, illumination is at oblique angles, and with axial symmetry, as indicated in Fig. 3.3(b). The zero-order beams reflect to the opposite side, beyond the objective lens. Light returning to the microscope is due only to diffraction and scattering from the substrate. A conventional darkfield microscope is not suitable for ISPI since it would block the exposing radiation, and the 360° illumination would be inefficiently utilized by the rectilinear diffraction gratings.

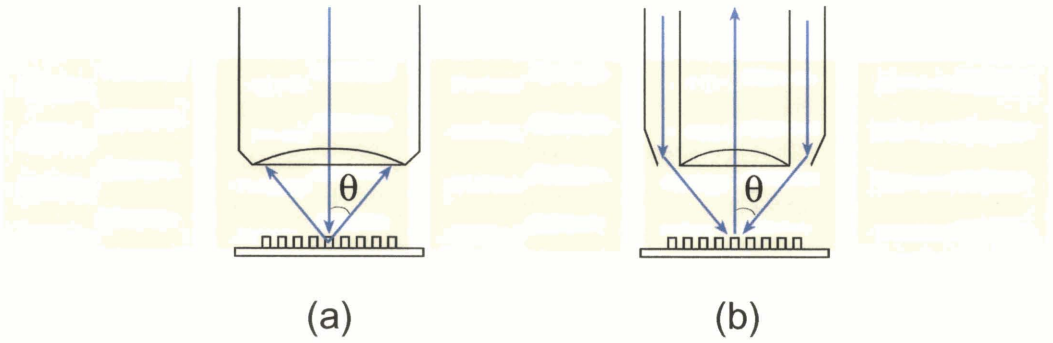


Fig. 3.3. (a) Schematic of a conventional brightfield microscope. Illumination is at normal incidence. First-order diffracted beams are collected by the objective lens. (b) Schematic of a darkfield microscope. Illumination is axially symmetric and at oblique incidence. First-order diffracted beams return to the objective lens. Zero-orders reflect to the sides, and escape capture by the lens. The darkfield microscope is suitable for viewing samples with arbitrary feature geometries. ISPI uses a specific feature geometry that allows specialized illumination and viewing angles.

We resolve the impasse by replacing the conventional axial symmetry of the illumination and imaging paths with symmetry perpendicular to the incident plane, effected with a periodic structure, as illustrated in Fig. 3.4.

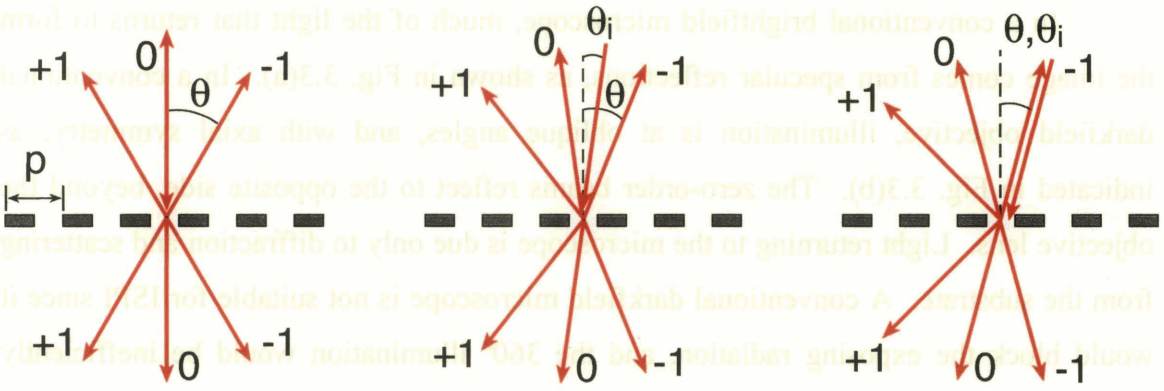


Fig. 3.4. Diffraction from a grating of period p at various angles of incidence. As the angle of incidence increases away from the normal, the angle of one first order diffraction decreases towards the normal. At a particular angle (called the Littrow angle) the first order beam returns along the same path as the incidence beam. The result is that a single optic can be used for both illumination and imaging at an oblique angle, allowing a convenient and practical implementation of an ISPI microscope. Imaging at the Littrow angle is important to feedback-controlled alignment and gap control at any time, including during exposure. Littrow imaging also acts as a filter to prevent any light from returning to the microscope except that which is designed to backdiffract at a particular angle, thus eliminating most degrading influences of defects or particles.

As the illumination angle increases away from the normal, backdiffracted orders rotate in the opposite direction; hence, we encounter a particular oblique angle at which the incident and backdiffracted beams are collinear, referred to as the Littrow angle. The Littrow angle θ_L is given by

$$\theta_L = \sin^{-1}\left(\frac{\lambda}{2p}\right) \tag{3.1}$$

where λ is the wavelength and p is the grating period. ISPI typically uses p = 1 μm and λ ~ 690 nm, resulting in $\theta_L \sim 20°$. Breaking axial symmetry allows the zero-order beams to reflect off to one side, thereby eliminating them from image formation. Light that returns to the ISPI microscope is light that is intentionally diffracted back to the microscope, resulting in higher efficiency and stronger signal-to-noise ratio. The Littrow angle is highly selective – every feature on the mask and substrate will scatter light; only gratings that have the designed periodicity will backdiffract at the microscope's viewing angle. There is no unique scattering angle, but there is a unique backdiffraction angle.

Although using the Littrow angle is preferable to using scattered light (see the Scattered Light Aligner in Section 2.1.2), a small amount of light returns to the ISPI microscope due to scattering from the miniscule edges of substrate topography that are perpendicular to the incident beam. The relative intensity of the scattered light depends on the surface roughness and radius of curvature at the corners of the features. In practice the intensity of the scattered light is negligible: it is estimated to be $\sim 10^{-5}$ less than the intensity of the ISPI fringes, and is usually below the detectivity threshold of the CCD cameras.

Since ISPI uses oblique-incidence illumination, light is received in the microscope objective lens only when there is a structure on the surface that diffracts light at the viewing angle of the microscope. The advantage of this diffractive dark-field imaging is that fringe quality can be improved and signal-to-noise ratio increased. In the

case of ISPI, this is true to a larger extent than in conventional dark-field microscopy, due to the highly selective design of the ISPI marks, which diffract light at a particular angle, and thereby eliminate all light returning to the microscope except that which is specifically designed to form the interference fringes.

The plane that contains the surface normal, the incident beam and the diffracted beams is referred to as the *incident plane*.

3.2.2 Diffraction Paths

In ISPI, we encode position information in the spatial phase of interference fringes. An interferometric moiré pattern is formed by interference of beams diffracted from periodic structures on the mask and substrate. Only two diffracted orders are required to form a fringe pattern, so we expect the interference fringes to be formed by interference of at least two orders, diffracted by both the mask and substrate gratings. The plane that contains the interfering beams is referred to as the *transverse plane*, and is perpendicular to the incident plane.

Consider diffraction of an incident beam from a periodic structure, as illustrated in Fig. 3.5. In the transverse plane, diffraction will occur from the substrate period p_1 at angles given by

$$\theta_1 = \sin^{-1}\left(\frac{m\lambda}{p_1}\right) \tag{3.2}$$

where λ is the wavelength of illumination and m is the order of diffraction.

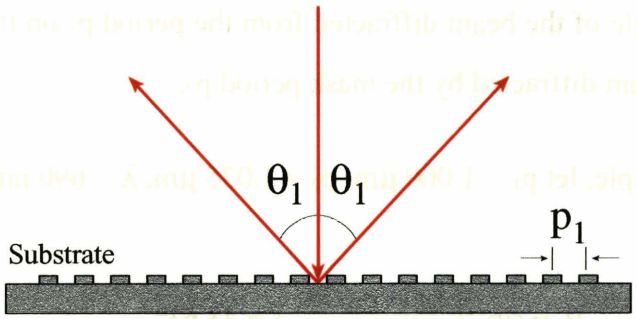


Fig. 3.5. Schematic of diffraction from a substrate grating in the transverse plane.

The beams that diffract from the substrate grating will diffract again at the mask grating, as illustrated in Fig. 3.6.

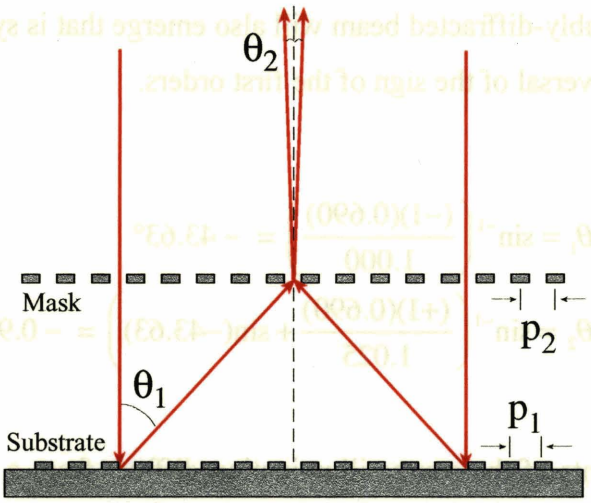


Fig. 3.6. Schematic of double diffraction from two different grating periods, resulting in interference moiré above the mask plane.

The angle of diffraction from the mask grating p_2 is given by

$$\theta_2 = \sin^{-1}\left(\frac{m\lambda}{p_2} + \sin\theta_1\right) \tag{3.3}$$

where θ_1 is the angle of the beam diffracted from the period p_1 on the substrate, and θ_2 is the angle of the beam diffracted by the mask period p_2.

As an example, let $p_1 = 1.000\ \mu m$, $p_2 = 1.025\ \mu m$, $\lambda = 690$ nm, and $m = +1$,

$$\theta_1 = \sin^{-1}\left(\frac{(+1)(0.690)}{1.000}\right) = 43.63^\circ$$
$$\theta_2 = \sin^{-1}\left(\frac{(-1)(0.690)}{1.025} + \sin 43.63\right) = 0.96^\circ \tag{3.4}$$

It is apparent that the double diffraction causes two large-angle diffractions to result in a small final diffraction angle.

A second doubly-diffracted beam will also emerge that is symmetric with the first, as a result of reversal of the sign of the first orders.

$$\theta_1 = \sin^{-1}\left(\frac{(-1)(0.690)}{1.000}\right) = -43.63^\circ$$
$$\theta_2 = \sin^{-1}\left(\frac{(+1)(0.690)}{1.025} + \sin(-43.63)\right) = -0.96^\circ \tag{3.5}$$

Thus, two components of the source illumination diffract from a mark on the substrate, meet at the mask, and diffract again[16]. The interference of these beams, diffracting at small angles, forms a fringe pattern with period P_F, given by

$$P_F = \frac{\lambda}{2\sin\theta_2} \tag{3.6}$$

[16] Each beam that contributes to image formation is diffracted in first order twice. A beam diffracts in first order from p_1 and then from p_2, or first from p_2 and then p_1. The same fringe period is produced regardless of the sequence of the +1 and -1 diffractions, resulting in identical periods in the interference pattern.

where the θ_2 is the final diffraction angle. Expressing θ_2 from Eq. 3.3 in terms of p_1 and p_2

$$\theta_2 = \sin^{-1}\left(\frac{m\lambda}{p_2} + \sin\theta_1\right) = \sin^{-1}\left(\frac{m\lambda}{p_2} + \frac{-m\lambda}{p_1}\right) \tag{3.7}$$

or. if we assume only the first order |m|=1

$$\theta_2 = \sin^{-1}\left(\frac{\lambda}{p_2} - \frac{\lambda}{p_1}\right) = \sin^{-1}\lambda\left(\frac{1}{p_2} - \frac{1}{p_1}\right) = \sin^{-1}\lambda\left(\frac{p_1 - p_2}{p_1 p_2}\right) \tag{3.8}$$

Substituting θ_2 into Eq. 3.6 yields

$$P_F = \frac{1}{2\left(\frac{p_1 - p_2}{p_1 p_2}\right)} = \frac{p_1 p_2}{2(p_1 - p_2)} \tag{3.9}$$

For arbitrary values of p_1 and p_2 the moiré fringe period is

$$P_F = \frac{p_1 p_2}{2|p_1 - p_2|} \tag{3.10}$$

For p_1 = 1.00 μm and p_2 = 1.025 μm this equation predicts a fringe period P_F = 20.5 μm. This expression for the fringe period is half the fringe period for geometric moiré, due to the removal of the zero order. Note that the contributions from the wavelength of illumination cancel, and the expression for P_F is achromatic. The wavelength can vary without affecting the fringe period, allowing the use of a wide spectral bandwidth of light: any wavelength of illumination produces the identical fringe period. In practice, we use a simulated broadband source, consisting of four laser wavelengths in the range $650 < \lambda < 690$ nm (described in Section 8.4). Due to the achromatic nature of the mark configuration, the laser wavelength is not a critical factor, which contributes to a highly

robust system. Unlike many of the alignment systems described in Chapter 2, laser frequency drift has no effect on the measured position.

The fringe pattern is also independent of gap. When the gap changes, the diffraction angles remain constant, but the path lengths and phase information in the subsequent diffracted beam vary, depending on the position of the grating from which the beam diffracts. The position on the grating, and hence the phase of a diffracted beam will vary with gap. However, since the grating period is constant, the gap-dependent phase change of one beam is exactly balanced by an equal and opposite phase change of the corresponding diffracted beam, as depicted in Fig. 3.7. When these two beams interfere, their phase difference, and hence the fringe pattern, is identical, regardless of the gap.

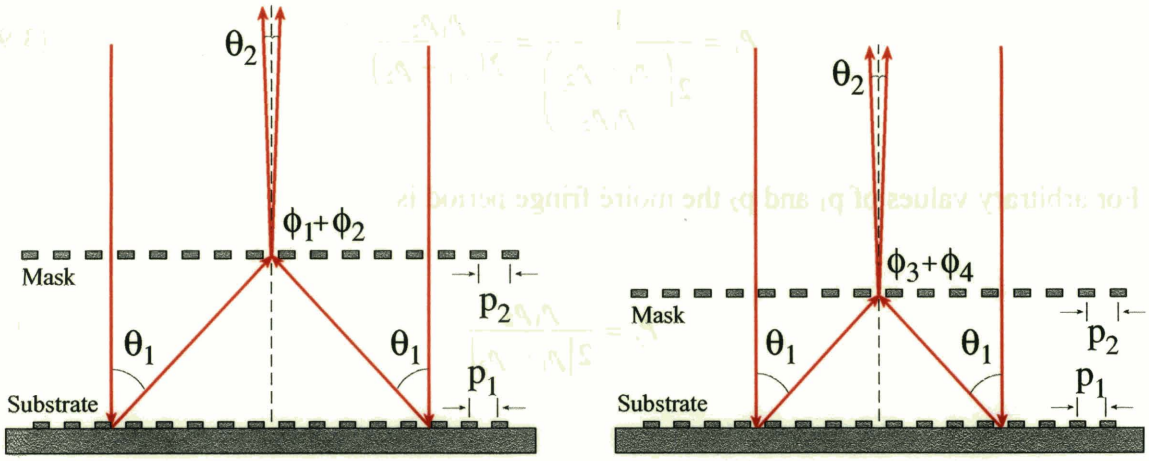


Fig. 3.7. Illustration of beams diffracted in the transverse plane at two different gaps (a) and (b). The phase of each beam diffracted at angle θ_1 will change, due to its sweep in position along the grating as the gap changes. However, the phase change of each beam will be equal and opposite, providing the grating periods are constant and the mask and substrate planes are parallel, resulting in an identical phase relation in the interfering beams.

As illustrated in Fig. 3.7, the phase of each beam diffracted by the substrate grating from either side is ϕ_1 and ϕ_2 at an initial gap. The phase of the beams diffracted at a slightly different gap are ϕ_3 and ϕ_4, where

$$\phi_3 = \phi_1 + \Delta\phi$$
$$\phi_4 = \phi_2 - \Delta\phi \tag{3.11}$$

resulting in

$$\phi_1 + \phi_2 = \phi_3 + \phi_4 \tag{3.12}$$

and identical fringe patterns, regardless of the gap.

Now consider a layer of resist covering an alignment mark on the substrate, as illustrated in Fig. 3.8.

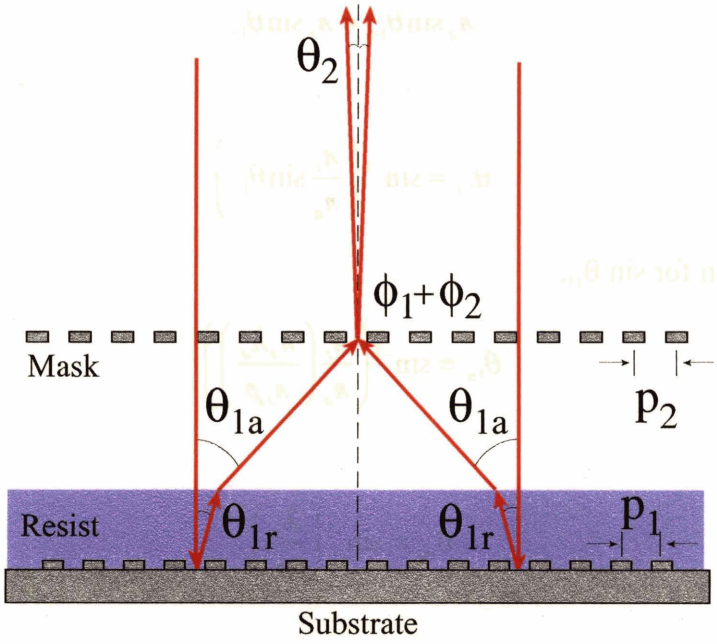


Fig. 3.8. Sketch of diffracted beams inside resist, taking note of the refraction of the diffracted beams. The wavelength change within the resist introduces a compensating factor, which allows cancellation of the index of refraction of resist in the expression for the refraction angle. The angle of a first-order diffracted beam above the resist is the same, with or without resist, although the optical path length changes with the presence of resist. The fringe frequency is unaffected by the resist because fringe formation depends only on the angles between the diffracted beams. The fringe phase is unaffected as a result of equal and opposite phase changes in the two paths.

The equations that describe the effect are

$$n_a\lambda_a = n_r\lambda_r \tag{3.13}$$

where the subscript "a" indicates air and the subscript "r" indicates resist, and

$$\theta_{1r} = \sin^{-1}\left(\frac{n_a\lambda_a}{n_r p}\right) \tag{3.14}$$

where θ_{1r} is the angle of the first order diffracted beam inside the resist. By Snell's law

$$n_a \sin\theta_{1a} = n_r \sin\theta_{1r} \tag{3.15}$$

Solving for θ_{1a},

$$\theta_{1a} = \sin^{-1}\left(\frac{n_r}{n_a}\sin\theta_{1r}\right) \tag{3.16}$$

Upon substitution for $\sin\theta_{1r}$,

$$\theta_{1a} = \sin^{-1}\left(\frac{n_r}{n_a}\left(\frac{n_a\lambda_a}{n_r p}\right)\right) \tag{3.17}$$

or

$$\theta_{1a} = \sin^{-1}\left(\frac{\lambda_a}{p}\right) \tag{3.18}$$

which is the equivalent to θ_1, the diffraction angle that would have occurred without any resist. Although the resist layer will cause a displacement of the diffracted beam, the angle of diffraction (hence the moiré fringe frequency) will remain the same, with or without the resist. Since the phase is changed by equal and opposite amounts in the beampaths within a constant-thickness resist, the alignment signal is unaffected by a resist layer.

3.2.3 Counterpropagating Fringe Motions

To determine motion, a set of interferometric moiré fringes must be compared to a reference. Camera pixels provide a crude reference, but accurate phase measurement depends on camera stability. A fixed reference on the substrate is another form of reference, as in the Moel, Moon, and Smith OAI scheme, but is subject to distortions from process etching, deposition, and resist layers. Such influences affect the reference marks differently than the gratings that produce the interferometric moiré fringes, due to the differences in mark size and spatial frequency. A fixed reference on the mask is more stable, but will still react differently than the fringes to variations in the index of refraction in the beampath.

As implied in Figs. 3.1 and 3.2, we eliminate differences between the fringes and their reference using a second set of fringes, adjacent to the first set, as the reference. The fringe sets have the same spatial frequency and appear identical to each other, as shown in Fig. 3.9.

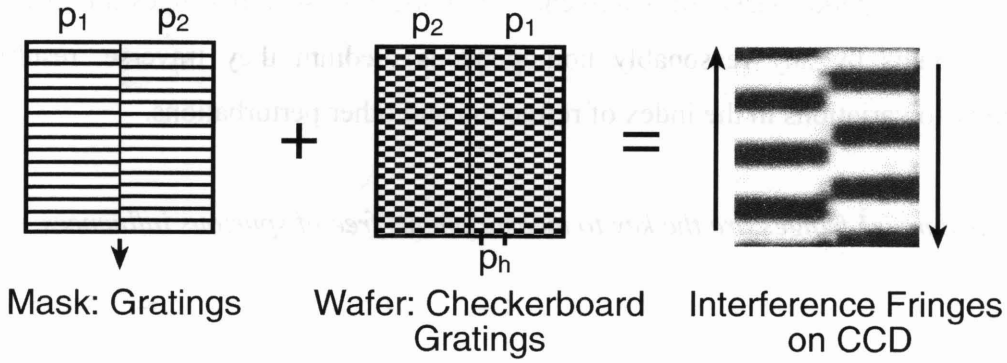


Fig. 3.9. Alignment marks and interference fringes resulting from inverted grating periods. Motion of the mask grating results in magnified, counterpropagating motions of the interference fringes. Period p_h directs the interference fringes towards the imaging optics at the Littrow angle.

One fringe set is produced by superposition of p_1 over p_2, and the other fringe set is produced by superposition of p_2 over p_1. The two fringe sets move in opposite directions due to the inversion of the grating periods. Counterpropagating fringe motion doubles sensitivity, since an identical phase condition is repeated in half the displacement required with a fixed reference.

We can define alignment as the position at which the spatial phases of the interference fringe patterns on the two sides of the CCD image plane match one another, or we can equally well define the aligned condition to be at any non-zero phase value, as would be advantageous in multiple exposures of the same mask with various known offsets[17]. (ISPI is probably unique in this respect.)

Encoding alignment in two matched fringe sets is a robust approach since the periodic structures have identical responses to etching and processing steps, by virtue of their nearly identical spatial frequencies. In addition, by encoding position in the relative phase of matched fringe sets, any perturbation caused by changes in the index of refraction in the beam paths will affect both fringe sets by equal amounts (assuming homogeneous media in the narrow ~400 μm beam width), thereby eliminating variations in the relative spatial-phase measurement. In short, both sets of fringes are affected by equal amounts by any reasonably homogeneous medium they traverse, resulting in immunity to variations in the index of refraction and other perturbations.

Symmetric moiré fringes are the key to measurements free of spurious influences.

[17] As used in constructing, for instance, a three-dimensional photonic band gap device

3.3 *ISPI Microscope: Low-NA, Littrow-Incidence, Dark-Field Imaging*

An ISPI microscope consists of 6x magnification optics with external single-mode fiber optic illumination, as illustrated in Fig. 3.10. The working distance is 110 mm and NA = 0.06. The microscope is oriented nearly parallel to the substrate, and a mirror in front of the objective lens steers the incident beam towards the mask and returns the interfering diffracted orders. The bisector of the incident and returning beams is the Littrow angle given by a p_h = 1.0 μm, and λ = 690 nm, or θ_L = 20.2°. Due to the large viewing angle, specularly reflected beams are discarded, and the only light that returns to the microscopes is light that is specifically backdiffracted by the alignment marks, yielding unusually high signal-to-noise ratio.

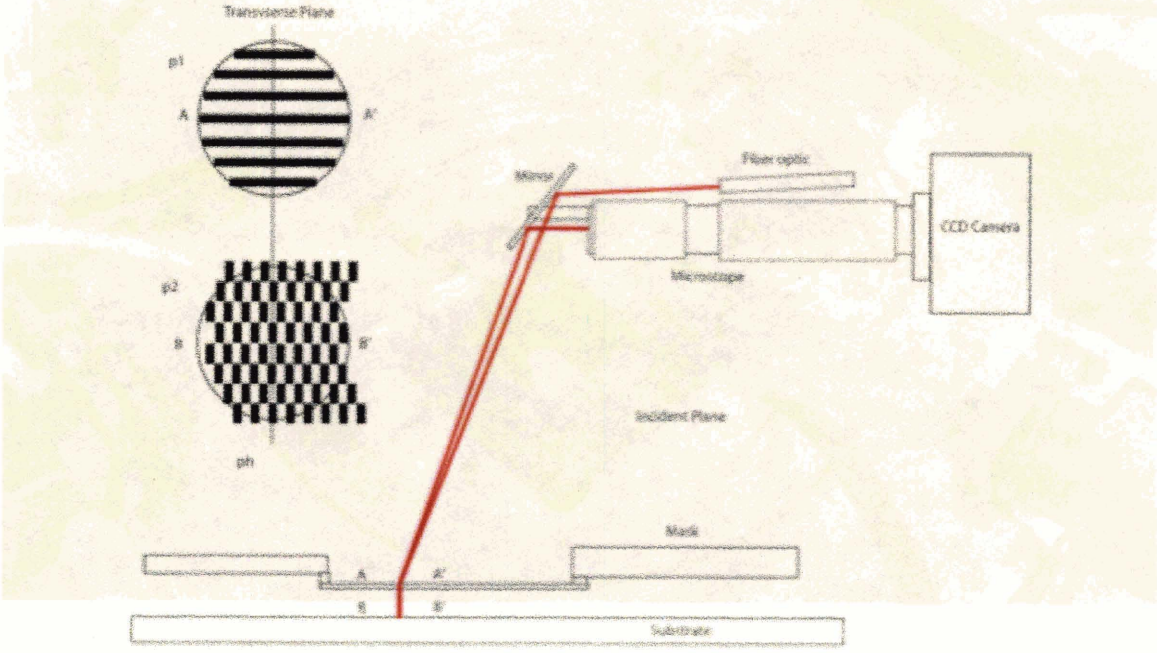


Fig. 3.10. Schematic of ISPI microscope. Laser light delivered via a single-mode fiber optic, reflects off a steering mirror at a 22° angle with respect to the mask normal. After diffraction from the mask and substrate marks, multiple beams interfere, return at an 18° angle, and are imaged by the 6x microscope onto a CCD camera. The working distance of the microscope is 110 mm. (The Littrow angle is 20°, the mean of 22° and 18°.)

The large viewing angle and long working distance result in the alignment mark position being ~35 mm in front of the steering mirror. In the MIT system, an exposure field typically contains a set of four die, 20 mm on each side, hence, the ISPI microscope is capable of viewing any point on the dies without obstructing the exposure. Fig. 3.11 shows three ISPI microscopes in typical viewing position. In this case two of the microscopes are imaging X and Y alignment marks at the center of the mask, and the third is imaging a Y alignment mark towards the edge of the exposure field.

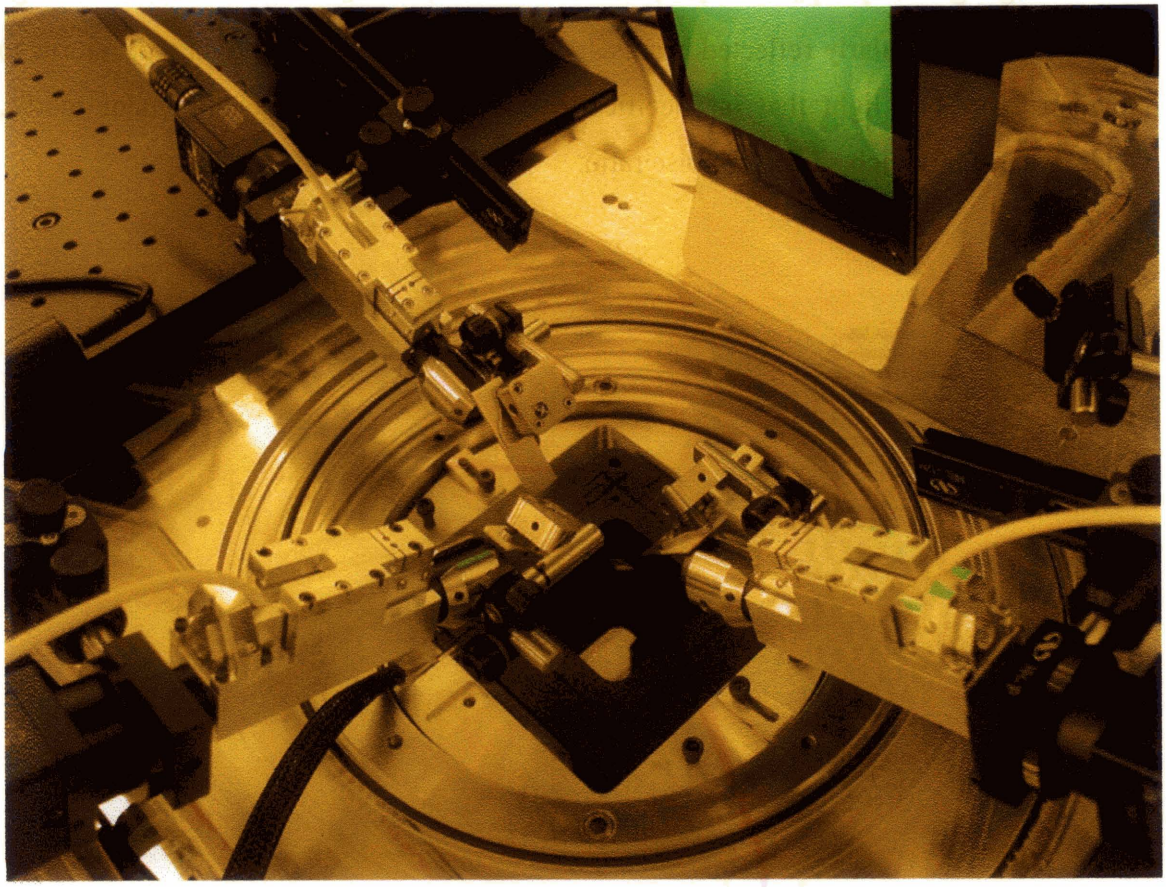


Fig. 3.11. Photograph of three ISPI microscopes viewing ISPI marks at oblique incidence. Illumination is brought in with single-mode fiber optics, collimated, and reflected down to the mask and substrate at a near-Littrow angle. Diffracted beams return at a similar angle, and are imaged onto a CCD camera by 6x microscopes at a working distance of 110 mm from the mask.

In most alignment systems, the low NA (0.06) of these microscopes would preclude fine alignment. In fact, the resolution of the microscopes is only 11 μm. Such low resolution becomes an advantage in ISPI, since it acts as a lowpass filter, rejecting high frequency image noise, such as particles or defects, and resulting in a clean, low-noise moiré signal. The moiré fringes have a period of 20 μm, or more. The NA of the microscopes was specifically chosen to filter the spatial frequencies of the grating marks, which would otherwise present an undesirable image component.

Another advantage of the microscopes is their depth of focus of ±150 μm. This is much larger than the ±1 μm focal depth of high-NA lenses typically used in alignment systems. ISPI microscopes can perform alignment over a wide range of gaps, and with insensitivity to small focal variations[18].

3.4 Displacement Magnification

Moiré displacement magnification is given by the expression

$$M = \frac{O_m C_p P_F}{P_{av}} \tag{3.19}$$

where O_m is the optical magnification, C_p is another factor of two from the counterpropagating fringes, P_F is the fringe period, and P_{av} is the average grating period. A plot of the moiré magnification is shown in Fig. 3.12. The fringe period P_F and average grating period P_{av} are given by

[18] A microscope's depth of focus determines the range of microscope Z positions over which geometric fiducial marks can be resolved. In marked distinction, interferometric fringe sets are limited by the spatial coherence of the illumination, and can remain in focus far beyond the nominal depth of focus of the microscope. The ISPI fringes are observed to remain in focus over >10 mm (observation limited by the microscope's Z stage travel).

$$P_F = \frac{p_1 p_2}{2|p_1 - p_2|} \tag{3.20}$$

and

$$P_{av} = \frac{2p_1 p_2}{p_1 + p_2} \tag{3.21}$$

where p_1 and p_2 are the grating periods, and are typically of the order of 1 μm.

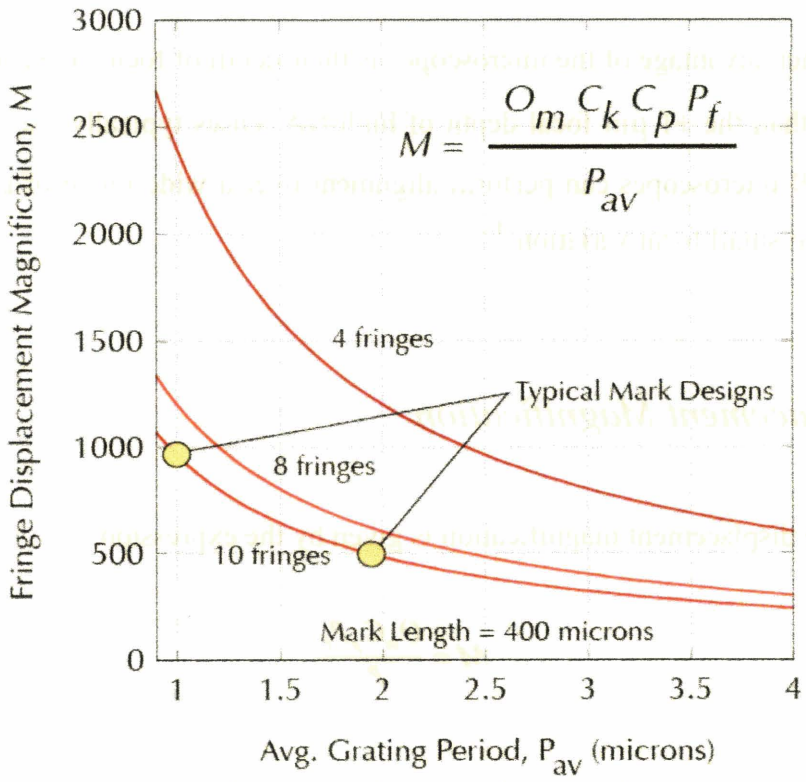


Fig. 3.12. Plot of the fringe displacement magnification as a function of average grating period. Typical magnification is >1000 with a 6x optical microscope. Displacement magnification is >10x greater than with a 100x optical microscope viewing geometric alignment marks.

Fig. 3.12 indicates that high fringe magnification is obtained with a small grating period and a limited number of fringe periods. The phase detection algorithm requires several fringes to achieve an accurate phase estimate, as discussed in the following section. With a 6x optical microscope, 10 fringes, and a 1 μm period, M ≈ 1000, or an

order of magnitude improvement over a 100x, high-NA microscope. High sensitivity is achieved with low-NA microscopes in combination with interferometric moiré fringes: an optical microscope with ~1/16 the magnification (6x vs. 100x) achieves 10x greater displacement magnification by using moiré fringes.

Typically we use grating periods p_1 = 2.000 and p_2 = 2.050 μm, producing fringe period P_F = 41.000 μm, or grating periods p_1 = 1.000 and p_2 = 1.025 μm, producing fringe period P_f = 20.500 μm. The mark length is typically 400 μm, resulting in 400μm/41μm = 9.75 fringes or 400μm/20.5μm = 19.5 fringes along the mark. For the ~2 μm gratings, the fringe magnification is M = 972, indicating that a displacement of the mask by 10 nm causes a fringe displacement of 9.72 μm on the CCD sensor (one CCD pixel is 7.4 μm). Sub-pixel sensitivity is obtained with a phase analysis algorithm, described in the next section.

3.5 Spatial Phase Analysis

ISPI uses frequency-domain analysis to extract a signal at a known frequency. The spatial phase relation between two sets of interference fringes is found by multiplication of the Fast Fourier Transform (FFT) of one fringe set with the complex conjugate of the FFT of the complementary fringe set[19]. The FFT gives a spectrum of frequencies in the image. A similar phase spectrum is found. We are interested only in the phase value corresponding to the frequency peak, i.e., the fringe frequency. Fig. 3.13 shows a block diagram of the algorithm.

In the algorithm, two regions of interest (ROI) are extracted from the captured image. The length of each region is set equal to N, the number of pixels perpendicular to the

[19] Since the complex conjugate changes the sign of the complex part in the exponential, the multiplication results in the phase difference of the two sets.

fringes. The width of each region is constrained by the boundary between the two fringe sets. In left and right ROI, the pixels are averaged along every row, forming one-dimensional signals, x[n] and y[n]. Each one-dimensional signal is multiplied by a windowing function to avoid introduction of spurious frequencies in the FFT, and zero padded. The FFT produces X[f] and Y[f], from which we find X*[f]/N and Y[f]/N. The product of X*[f]/N and Y[f]/N yields exponentials with the difference of the phase at every frequency, seen in plots of the magnitude and phase in Fig. 3.13, where the phase difference of the two fringe sets is given by the value of the phase spectrum at the index corresponding to the frequency peak.

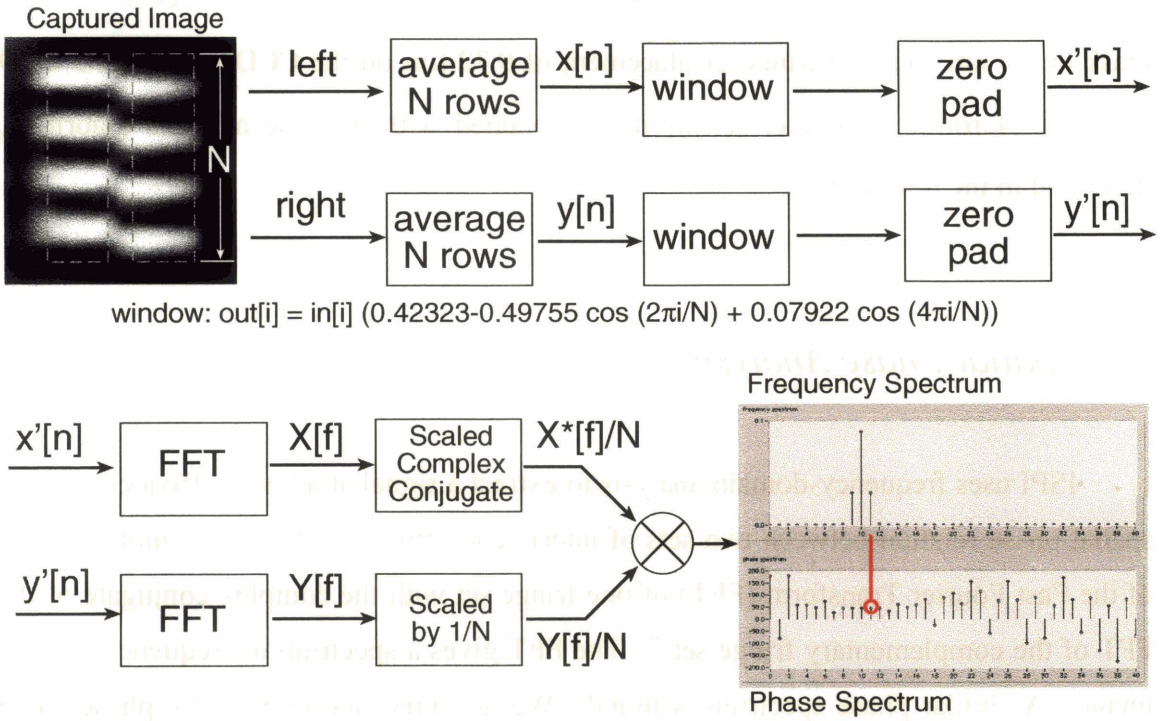


Fig. 3.13. Block diagram of the basic phase analysis algorithm used to determine the phase difference between two sets of ISPI fringes.

This phase value is converted into a displacement with the expression

$$\Delta x = \frac{\phi P_{av}}{8\pi} \tag{3.22}$$

where P_{av} was given by Eq. 3.21.

Including the effects of noise, the phase can be measured with an efficient estimator, such as the Discrete Fourier Transform used in the Fast Fourier Transform. The limit on phase estimation is given by the Cramer-Rao bound[20], indicating that the standard deviation of the phase estimate is

$$\sigma_\phi = \sqrt{\frac{2(N-1)}{(SNR)(N)(N+1)}} \approx \sqrt{\frac{2}{(SNR)(N)}} \tag{3.23}$$

for large N, where N is the number of samples, and SNR is the signal-to-noise ratio. The SNR is given by

$$SNR = \frac{A^2}{\sigma^2} \tag{3.24}$$

where A is the amplitude of the ISPI fringes and σ is standard deviation of the noise.

If the standard deviation of the phase is given by σ_ϕ, then the standard deviation of the position measurement is given by

$$\sigma_x = \frac{\sigma_\phi P_{av}}{8\pi} \tag{3.25}$$

In ISPI we obtain high SNR: the dynamic range of the camera is 0 to 255, and the intensity of the lasers is adjusted so that the amplitude of the ISPI fringes spans nearly the entire range (typically the fringe amplitude is ~250). Noise is low, with $\sigma \sim 5$, yielding SNR = 2500. The number of pixels that are used along one set of fringes is typically N = 300. The standard deviation of the phase estimate is therefore

[20] The Cramer-Rao bound determines the minimum variance of a physical variable in the presence of noise [3.1].

$$\sigma_\phi = \sqrt{\frac{2(300-1)}{(2500)(300)(300+1)}} = 0.001628 \tag{3.26}$$

Therefore we expect the standard deviation in the position measurement can be

$$\sigma_x = \frac{(0.001628)(1)}{8\pi} = 0.065\text{nm} \tag{3.27}$$

Note that any phase disparity is measured with equal accuracy. Because ISPI can measure any phase discontinuity, not merely zero-phase, to our knowledge it is the only scheme that can measure a displacement from a nominal alignment point - a capability of significance in many current and future applications.

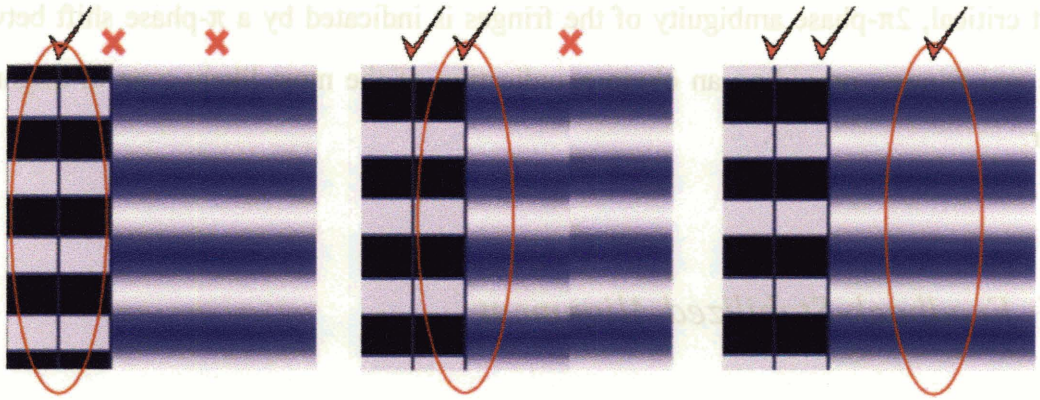
Several other advantages result from this algorithm. One of the advantages is that it uses all of the relevant information in the image, distilled into the spatial phase discontinuity. This method is more robust than an edge-detecting approach since all pixels of the image contribute to the measurement, not just those in the vicinity of edges.

Another advantage is that the spatial-phase algorithm filters out contributions from spurious spatial frequencies. Alignment is encoded in a specific, known spatial frequency. Particles or defects on the alignment marks may contain spatial frequencies that are not removed by lowpass optical spatial filtering. However, these types of defects will not normally be periodic at the same frequency as the fringes. Since analysis of the spatial-phase difference is at a single, known spatial frequency, contributions of spurious spatial frequencies are excluded. In the end, we extract from the image only spatial frequencies that correspond to the predetermined moiré frequency.

3.6 *Wide-Dynamic-Range Phase Measurement*

Any scheme that encodes information in the relative position of periodic patterns encounters positions at which the fringes appear indistinguishable from their starting position. Some means of resolving phase ambiguity, i.e., distinguishing between identical phases separated by 2π intervals, must be employed.

In ISPI, we take advantage of the ability to simultaneously analyze both *geometric* and *interferometric* features in the same image. Adjacent to the moiré fringe sets, we place small-area gratings, with their k-vectors parallel to both the incident plane and the moiré fringes. Each small-area grating backdiffracts directly to the microscope and is imaged as a bright spot. A set of such small-area gratings is referred to as a *bar array*. The period of the bar array is designed to match the spatial period of the moiré fringes. One bar array is on the mask, and another is on the substrate, in a non-overlapping position. Unambiguous alignment is achieved through a three-stage process, as illustrated in Fig. 3.14.



(a) (b) (c)

Fig. 3.14. Schematic of unambiguous phase alignment. A set of gratings with periodic interruptions (left) backdiffract bars of light to the ISPI microscope, adjacent to the moiré interference fringes (right). Unambiguous alignment consists of three steps: (a) Coarse alignment: Lining up the spatial phase of bars indicate coarse alignment from >30 mm to 30 nm. (b) Spatial phase between one set of bars and one set of fringes indicates intermediate alignment. (c) Fine alignment is determined by matching the spatial phase between two sets of interference fringes. The same phase algorithm is used in each case, with the exception of a scaling factor involved in the conversion of phase to displacement.

In the first stage, the relative spatial phase of the mask and substrate bar arrays indicates coarse alignment. In the second stage, the relative spatial phase between aligned bars and one moiré fringe set determines intermediate alignment. In the third stage, the phase between the moiré fringes indicates fine alignment. The same spatial phase algorithm described in the previous section is used for each stage, with the appropriate phase-displacement conversion factor in each case[21]. A simple modification to the algorithm that analyzes the bar-bar phase (detection of unmatched bars on either mask on substrate, indicating displacement by more than one bar period) increases coarse position detection to the extent of the field of view of the microscope. In this way, spatial-phase ambiguity is eliminated and the dynamic range extends from <0.5 nm to >500 μm, or a factor of 10^6.

Although there is significant overlap between the detectivity of the bar-bar phase and the phase ambiguity of the fringes, the intermediate step involving bar-fringe phase provides effective assurance of the desired position: the bar-fringe phase is π-phase-shifted for every 2π-phase shift of the fringes, as illustrated in Fig. 3.14(a). Thus the most critical, 2π-phase ambiguity of the fringes is indicated by a π-phase shift between bars and fringes, providing an obvious indication of the most likely type of alignment error.

3.7 Feedback-Stabilized Alignment

An example of feedback control, using ISPI measurements to control the mask position, is shown in Fig. 3.15. Without feedback, the mask position was observed to drift by >150 nm during six hours. Feedback from the interferometric signal detected by

[21] A bar array can be considered to be a "fringe set" with zero displacement magnification. The displacement of the bars is given by $\Delta x_B = (\phi\ P_F)/(4\pi)$, where ϕ is the bar-bar phase and $P_F = 41$ μm. Minimum detectivity of the bar array is ~30 nm.

an ISPI microscope was activated during the subsequent six hours, maintaining the mask-substrate position with a zero mean and σ = 1.4 nm. Details of the piezos and alignment apparatus used in this experiment are described in Chapter 8, and feedback-controlled overlay in Chapter 9.

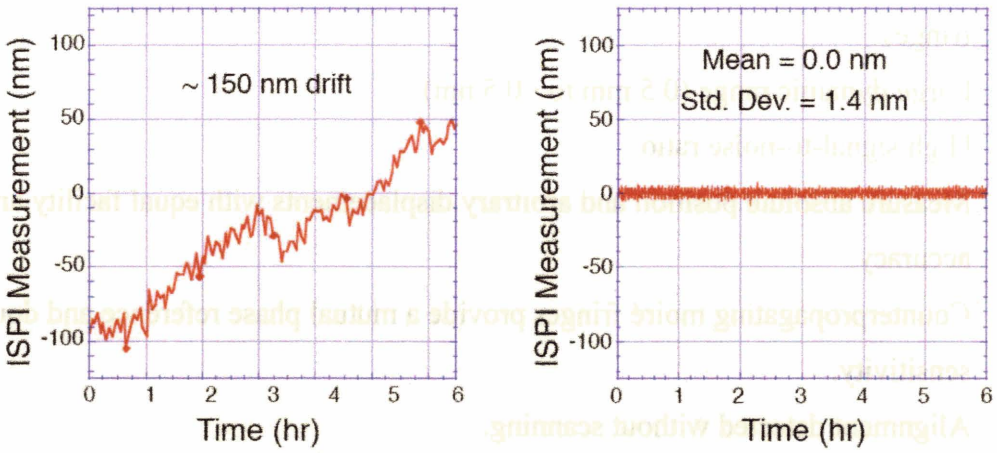


Fig. 3.15. (a) Drift and vibration in relative mask-substrate position detected with an ISPI microscope during six hours. (b) Six-hour alignment data with continuous closed-loop operation. Feedback-stabilized ISPI alignment nullifies the effects of thermal/mechanical drift to zero mean and σ = 1.4 nm.

3.8 *Summary*

Interferometric Spatial Phase Imaging (ISPI) combines a form of dark-field optical microscopy with interferometric moiré fringes, formed between specialized marks on the mask and substrate. The unique aspects of ISPI are

- Alignment is encoded in the spatial phase relation of two matched sets of moiré fringes.
- Large dynamic range (0.5 mm to <0.5 nm)
- High signal-to-noise ratio.
- Measure absolute position and arbitrary displacements with equal facility and accuracy.
- Counterpropagating moiré fringes provide a mutual phase reference and double sensitivity.
- Alignment detected without scanning.
- Feedback-stabilization of alignment during lithographic exposure.
- Alignment signal insensitive to fluctuations in illumination wavelength or intensity.
- Alignment signal insensitive to mask-substrate gap.
- Alignment signal insensitive to resist layers on substrate (described in Chapter 7).
- Alignment signal insensitive to changes in the index of refraction in the beampath.
- Image formation at a large oblique angle (20°) and long working distance (110 mm).
- Spatial filtering by low-NA microscope.
- Geometric filtering by oblique viewing angle.
- Elimination of zero-order diffraction components.
- Electronic filtering by phase measurement at a single spatial frequency.
- Broadband illumination eliminates thin-film interference effects.

- Illumination wavelengths do not expose resist.
- Detect alignment and rotation in the same image (and gap - Chapters 4 and 5)
- Low-cost optics

These advantages make ISPI efficacious in a wide variety of proximity lithography applications. In Table 3.1 we summarize advantages of ISPI and compare them with heterodyne interferometry and four other alignment methods that use some form of interferometry:

	Heterodyne Interferometer	Flanders et al.	TSA	TTM	OAI	ISPI
Detect Displacements <1 nm	YES	NO	NO	NO	NO	YES
Directly Measures Mask-Substrate Position	NO	YES	YES	YES	YES	YES
Absolute or Relative Displacement Measurement	Relative Displacement	Absolute Position (after scan)	Absolute Position (after scan)	Absolute Position (after scan)	Absolute Position	Absolute Position
No Scanning Required for Measurement	NO	YES	Beam Hop	NO	YES	YES
Feedback Stabilization During Exposure	YES	NO	NO	NO	NO	YES
Insensitive to Index of Refraction in Beampaths	NO	NO	NO	NO	NO	YES
Insensitive to Layers of Resist Covering Substrate	NA	NO	NO	NO	NO	YES
Displacement Encoded in Relative Changes of Two Matched Signals	YES Two Frequencies	NO	YES	NO	NO	YES Two or More Spatial Phases
Detect Mask-Substrate Alignment	NA	YES	YES	YES	YES	YES
Insensitive to Gap	NA	NO	NO	NO	NO	YES
Detect Mask-Substrate Rotation with One Mark	NA	NO	NO	NO	YES	YES
Multi-Axis Measurements at a Single Point	NO	NO	NO	NO	NO	YES

Table 3.1. Comparison of features in interferometric alignment methods

4 Interferometric Gap Detection Methods

Gap: the separation or distance between two parallel surfaces.

This chapter will describe several means of optical gap detection that can be used in conjunction with ISPI aligning. Large gaps (between 300 and 30 μm) are detected with ±1 μm sensitivity from the geometric relation of backdiffracted and reflected beams originating from a single grating on the mask. Using adjacent p_1 and p_2 gratings on the mask, *interferometric* sensitivity (<1 nm) at large gaps is achieved with fringes that are similar to ISPI alignment fringes in appearance, but are sensitive only to gap. Below 30 μm gaps, an inclined Fabry-Perot interferometer determines gap by detecting the phase between intensity oscillations in scans of gap with different wavelengths of illumination.

4.1 Background of Gap Detection

A number of schemes for measuring the mask-substrate gap have been described in the literature. Capacitive sensors are the most common means of gap detection, and are used in most commercial nanopositioning systems to detect gaps of <100 μm. Detectivity scales with the gap range – for a gap range of 10 μm, detectivity can be <1 nm [4.1]. In a mask-substrate aligner, capacitive sensors are typically used to measure gap at three points around the periphery of the mask stage. Since the capacitive sensors measure the mask-substrate gap indirectly, calibration of the capacitive measurements to the actual mask-substrate gap is required. Errors associated with this scheme include thermal drift, and variations between dimensions of individual masks, and stress-induced deformation by the mask stage. With concerted efforts, the total gap detection error can be reduced to the ~1 μm level.

To avoid one of the secondary gap references, capacitive sensors can be embedded within part of the substrate stage, adjacent to the substrate, and scanned underneath the mask. Some problems with this are the potential mismatch between the capacitive sensing plane and the substrate plane, the time required to move the sensor and measure gap at multiple points, and perturbations in the capacitive measurements caused by patterned regions on the mask.

Aside from the secondary referencing problems, the intrinsic limitations of capacitive sensors lie in the tradeoff between range and resolution. As we can see from the capacitance relation, the capacitance falls off linearly with gap, G, between the plates:

$$C = \frac{\varepsilon_r \varepsilon_o A_p}{G} \tag{4.1}$$

where ε_r is the relative permittivity of the material between the plates, ε_o is the permittivity of vacuum, and A_p is the area of the plates.

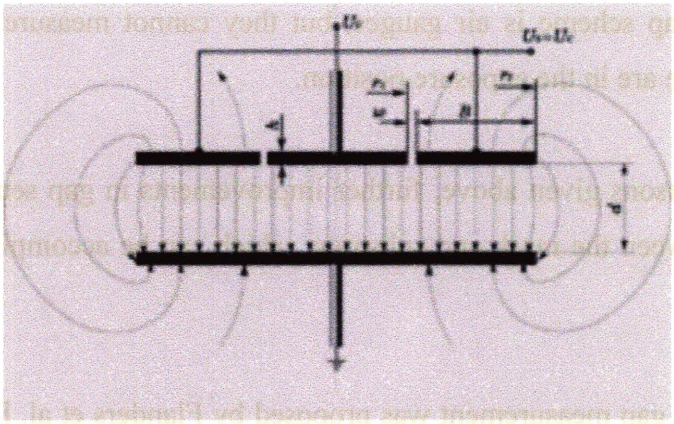


Fig. 4.1 Schematic of a parallel-plate capacitor.

At large gaps, electronic noise dominates, and limits gap range. At small gaps, noise is reduced, but sensitivity to non-parallelism of the capacitor plates is increased. In combination, these effects serve to limit the useful operating range.

In some applications, such as scanning probes, capacitive sensors are used to find the probe position in six axes, requiring six separate sensors. Since the range of capacitive sensors is very small compared to the diameter of the plates and the points of measurement are separated from each other and the scanning probe by a few centimeters, thermal expansion of the intervening lengths of material can degrade measurement accuracy.

Furthermore, in a multi-axis capacitor arrangement, the area of capacitor plate overlap will come into play during orthogonal motions; as one axis in an XY stage is moved, the area of plate overlap in the orthogonal axis will decrease, and may cause an undesirable variation in capacitance along the orthogonal axis.

Finally, capacitance measurements are generally somewhat awkward to implement, and do not apply to all types of substrates.

Another gap scheme is air gauges, but they cannot measure the gap when the mask and substrate are in the exposure position.

For the reasons given above, further improvements in gap sensing require direct measurement between the mask and substrate, which can be accomplished using optical methods.

An optical gap measurement was proposed by Flanders et al. [4.2], shown in Fig. 4.2, based on the interference of beams reflected from a HeNe laser *focused onto the* mask at an oblique angle. The diverging reflections from the mask and substrate interfered, and the interference fringes were detected by a linear photodetector array. Gap was determined by the spatial frequency of the interference pattern, which increased linearly with gap to first order, with a small gap-dependent deviation from linearity. The range of gap detection was between 25 and 120 μm, and the gap uncertainty was a percentage of gap, ranging between 120 nm and 600 nm.

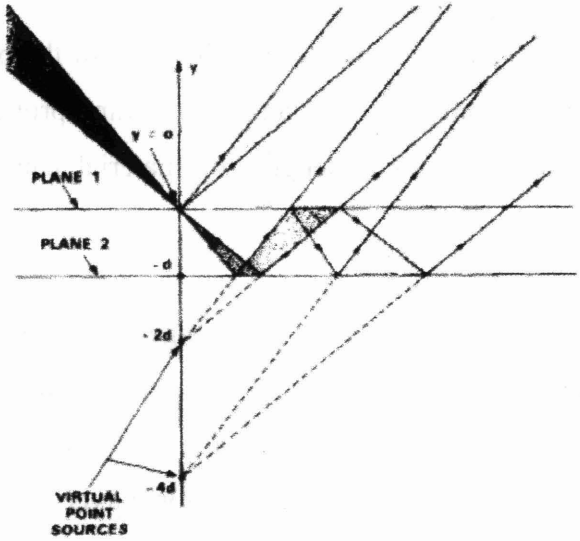


Fig. 4.2. Schematic of Flanders gap scheme. A focused laser (incident from the left) reflects from the mask and substrate at a 45° angle. Patterns characterisic of the gap are formed by interference between beams reflected from the mask and the substrate (planes 1 and 2, respectively). Virtual point sources for multiply-reflected beams are at even-integer multiples of the gap.

In proximity lithography, the exposure gap (i.e., <5 μm) is below the minimum detection range in the Flanders scheme (25 μm). At such small gaps the angular disparity between reflected beams is too small to produce measurable interference fringes. Also, the finite thickness of the mask membrane causes multiple spurious internal reflections, which corrupt the measurement signal. The Flanders scheme would not integrate well with the ISPI aligning scheme, since it requires diverging illumination. In addition the geometry of the Flanders scheme is incompatible with ISPI alignment microscopes since the Flanders scheme requires a different illumination angle and a separate sensor that must be on the opposite side of the mask from the illumination. In sum, the Flanders scheme is not compatible with ISPI, nor does it have the desired gap range or sensitivity.

Suzuki and Une [4.3] also described a gapping scheme based on the interference of diffracted beams, as shown in Fig. 4.3. In this scheme a beat signal I_{WZ} depended on interference of a first-order beam and a second-order beam, diffracted from a grating. Each order had a different diffraction angle, which also depended upon the laser frequency: a laser beam of frequency f_1 was incident at an angle θ_1, and another beam with frequency f_2 was incident at an angle θ_3. The two orders where arranged to diffract at the same angle, θ_2. The gap was linearly related to the phase of the beat in I_{WZ}. Sensitivity was ~32 nm. This scheme was sensitive to both lateral (Fig. 4.3(a)) and vertical (Fig. 4.3(b)) displacements. Displacement had to be measured simultaneously, as described in Section 2.4.1, and subtracted from the total signal to determine the gap. Although this scheme could detect small gaps used in proximity exposures, the inseparability of the gap signal from alignment was less than optimal.

The factors described above point to the advantages of using an optical method, but clearly indicate the difficulty of inventing a satisfactory gapping scheme.

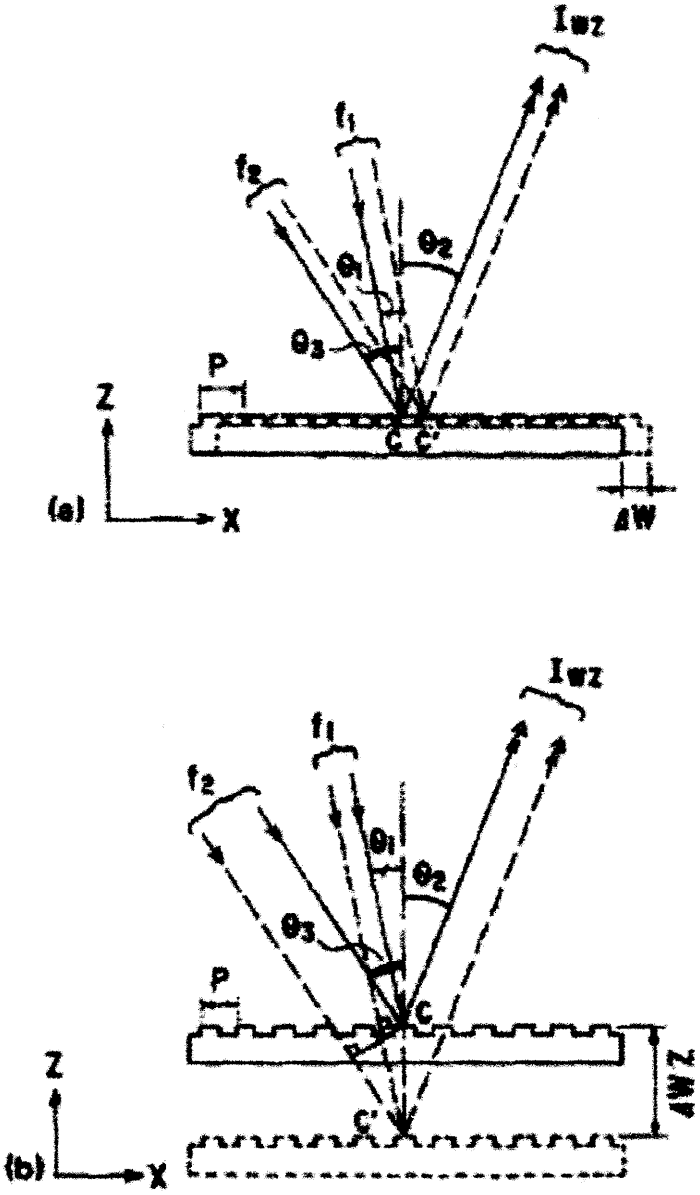


Fig. 4.3. Diagram of the gap detection method of Suzuki and Une. A beat signal depends on interference of a first order beam diffracted from a laser beam of frequency f_1, and a second order diffracted laser beam of frequency f_2. The gap is linearly related to the phase of the beat. Sensistivity is ~32 nm. This scheme is sensitive to both lateral (a) and vertical (b) displacements. Displacement must be measured simultaneously, as described in Section 2.4.1, and subtracted from the total signal to determine the gap.

4.2 Geometric Coarse Gap Detection

In any proximity lithography system, two objects need to be brought into propinquity prior to exposure. Control must be maintained over the gap while it is reduced from several hundred microns to a small exposure gap (<5 μm), and gap non-uniformity must be eliminated to ensure exposure consistency and avoid destructive consequences to the mask. For such mask leveling and initial gap descent control we devised a simple optical gapping method that is compatible with ISPI alignment.

The idea is shown in Fig. 4.4, where a single wavelength of light illuminates a small-area grating on the mask at an oblique angle. This small-area grating is typically 5x20 μm, with a 1 μm period, the same period used by the alignment marks to backdiffract light at the Littrow angle. The incident beam produces two rectangular spots imaged by an ISPI microscope. One spot is due to direct backdiffraction from the grating and the other spot results from a symmetric forward diffraction, that reflects off the substrate at the same angle as the backdiffracted beam. The beam that reflects off the substrate is separated from the backdiffracted beam by an amount, ΔS, given by

$$\Delta S = (2G\tan\theta)\cos\theta = 2G\sin\theta \tag{4.2}$$

where G is the gap, θ is the angle of illumination, and the cosθ term is due to image foreshortening in the incident plane direction. A peak detection algorithm measures the separation of the beams. Absolute separation is calibrated from the known microscope magnification and CCD pixel size. Absolute gap detectivity from these marks is <1 μm[22], ranging from the maximum focal depth of the microscope (300 μm) down to the

[22] For comparison, this is similar to the gap resolution attained in the Karl Suss system (Section 2.1.1) using the focal difference of high-NA optical microscopes.

gap at which the beams begin to overlap (at a 30 μm gap[23]). The large focal length of a low-NA ISPI microscope resolves both spots across a 300 μm gap range.

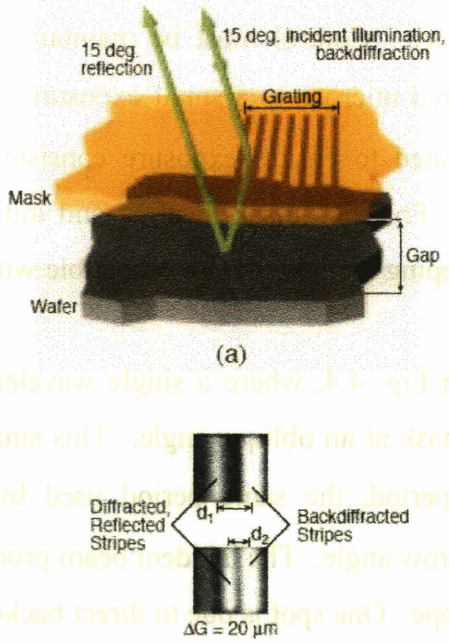


Fig. 4.4. (a) Schematic of diffracting beams. (b) Images taken with an ISPI microscope of the returning beams at two different mask-substrate gaps.

Gaps are measured simultaneously at three points with three ISPI microscopes, and using these measurements, the mask is automatically leveled with respect to the substrate by minimizing the differences in gap between measurement points, as illustrated in Figs. 4.5 and 4.6. Assuming a separation of 10 mm between points, the mask can be leveled to <0.1 arcsec. The mask is then brought within 30 μm of the substrate using an average of the three gapping measurements. For many applications, this relatively coarse, sub-micron gap detection method may be sufficient, although for nanoimprint lithography and x-ray lithography, gaps below 30 μm must be detected by other means.

[23] The gap at which overlap occurs depends in part upon the grating width.

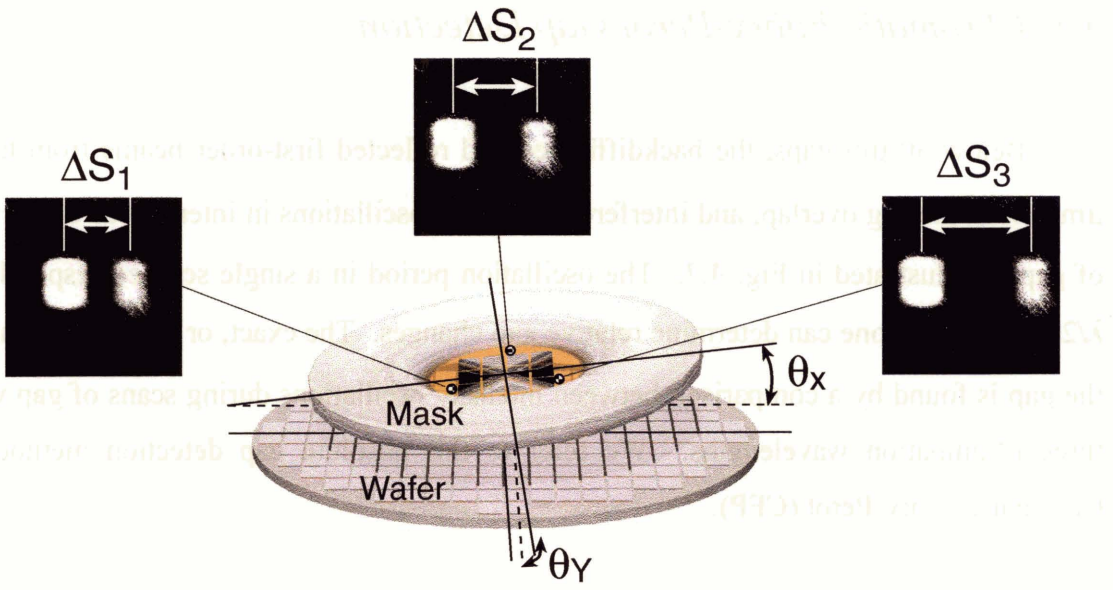


Fig. 4.5. Illustration of three-point measurement of gap using the diffracted beam separation. Mask tilt is indicated by differential separation.

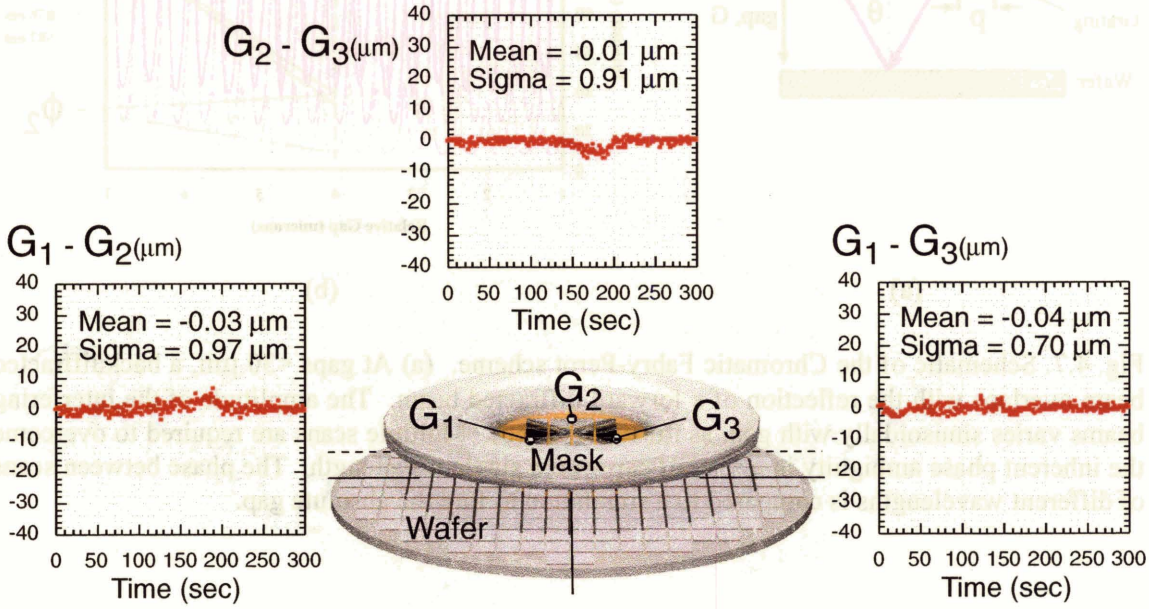


Fig. 4.6. Measurements taken after leveling and with feedback control of gap using the diffracted bar method.

4.3 *Chromatic Fabry-Perot Gap Detection*

Below 30 μm gaps, the backdiffracted and reflected first-order beams from the 1 μm-period grating overlap, and interfere, producing oscillations in intensity as a function of gap, as illustrated in Fig. 4.7. The oscillation period in a single scan corresponds to λ/2, from which one can determine relative gap changes. The exact, or absolute, value of the gap is found by a comparison between intensity oscillations during scans of gap with three illumination wavelengths. We refer to this absolute gap detection method as Chromatic Fabry-Perot (CFP).

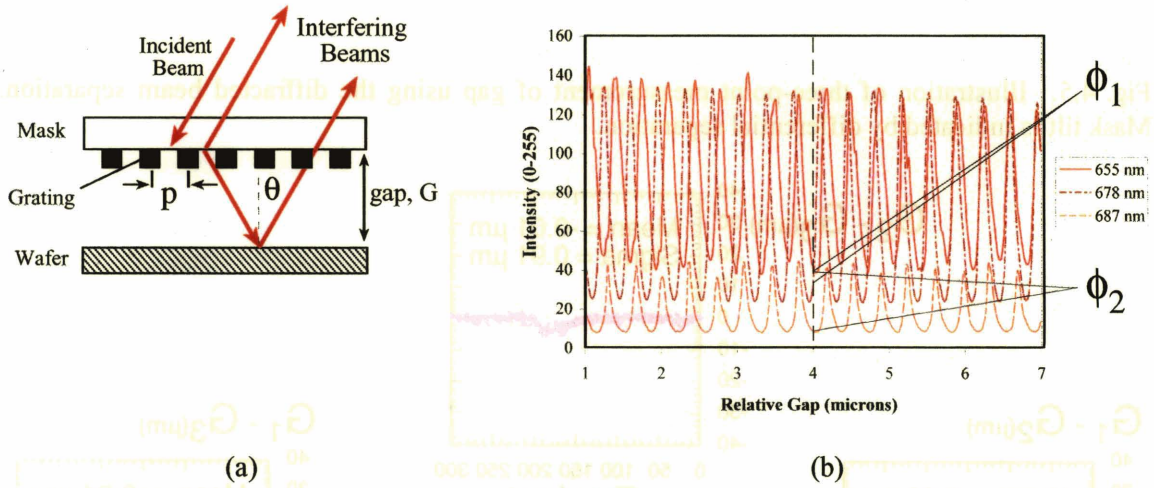


Fig. 4.7. Schematic of the Chromatic Fabry-Perot scheme. (a) At gaps <30 μm, a backdiffracted beam overlaps with the reflection of a forward diffracted beam. The amplitude of the interfering beams varies sinusoidally with gap, as illustrated in (b). Multiple scans are required to overcome the inherent phase ambiguity in a scan of gap with a single wavelength. The phase between scans of different wavelengths is compared to a simulation to find the absolute gap.

The CFP procedure is to: a) scan over a range of gaps while recording the intensity, using one scan over the same range of gaps for each of three different wavelengths, b) find the phase difference ϕ_1 between a pair of scans at the midpoint of

the scan range, c) find another phase difference ϕ_2 between a pair of scans with different wavelengths, d) compare ϕ_1 and ϕ_2 to phase differences found in simulated scans, starting from zero gap, and e) find the absolute gap by the point in the simulation where there is a simultaneous match between the simulated and experimentally determined ϕ_1 and ϕ_2. Simulated scans, and plots of potential phase matches are illustrated in Fig. 4.8.

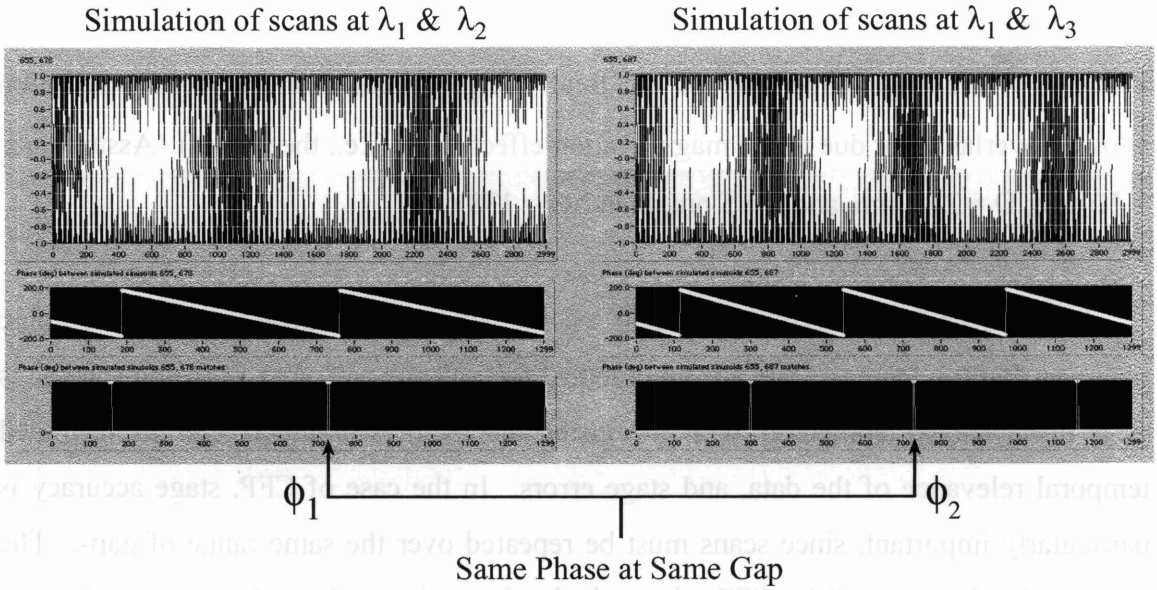


Fig. 4.8. (a) Simulation of scans at wavelengths λ_1 and λ_2, from 0 to 30 μm gaps. Beats are indicative of the relative phase of the two intensity plots. (b) Simulation of CFP scans at wavelengths λ_1 and λ_3. (c), (d) Plots of the relative phase between simulated scans. (e), (f) Plots of the correlation between experimental and simulated scans. At several points there are possible matching points of the experimental phase and simulated phase. A unique matching point indicates the absolute gap.

In the sub-30 μm interferometric regime, the CFP scheme increases sensitivity considerably. One period of a CFP intensity oscillation is equal to $\lambda/2$, corresponding to ~350 nm. Gap sensitivity is approximately $\lambda/10$, corresponding to ~35 nm. In terms of sensitivity, the CFP scheme meets the requirements of many forms of proximity lithography. In the case of collimated x-ray lithography of sub-100 nm features, the x-ray

diffraction that occurs in the gap between mask and substrate can lead to loss of control of feature sizes, unless the gap is known and controlled to <800 nm [4.4]. In the case of a non-collimated x-ray source, such as we use for device exposures in Chapter 9, tolerance to gap deviations is reduced because of the added effect of image magnification. For example, with a point source located at a distance D from a mask, the allowed uncertainty in gap ΔG, is given by

$$\Delta G = \varepsilon \frac{D}{L} \tag{4.3}$$

where L is the edge dimension of the field being replicated and ε is the maximum allowed overlay error due to the magnification effect alone (i.e., the runout). Assuming L = 10 mm, D = 275 mm, and ε = 5 nm, then ΔG = 138 nm.

The 35 nm detectivity of the CFP scheme is more than sufficient for lithographic purposes, but the scheme suffers from practical inconveniences. Any scanning scheme is at a disadvantage compared with an instantaneous measurement, due to degraded temporal relevance of the data, and stage errors. In the case of CFP, stage accuracy is particularly important, since scans must be repeated over the same range of gaps. The primary disadvantage of the CFP scheme is that it requires at least three scans of gap to determine the absolute gap at the midpoint of the scan, and obtain a single measurement.

4.4 Long-Range Interferometric Gap Detection

Geometric gapping marks operate over a 30-300 μm range, but their resolution is limited to ~1 μm, as indicated in Section 4.2. In certain situations, higher sensitivity, and a larger gap range are required. For such situations we encode gap in the frequency and phase of a set of fringes, produced by interference between two gratings of different periods on a mask. The gratings have periods p_1 and p_2, just as in ISPI, but are arranged

with the grating k-vectors along a single axis in the incident plane. so that at a finite gap the reflected first orders from one grating overlaps the backdiffracted beams of the other grating, as shown in Fig. 4.9. (In ISPI the k-vectors of the p_1 and p_2 gratings are parallel to each other and are in the transverse plane). In this Long-Range Interferometric Gapping (LRIG) method, as in the other ISPI-compatible gapping methods, no marks are required on the substrate.

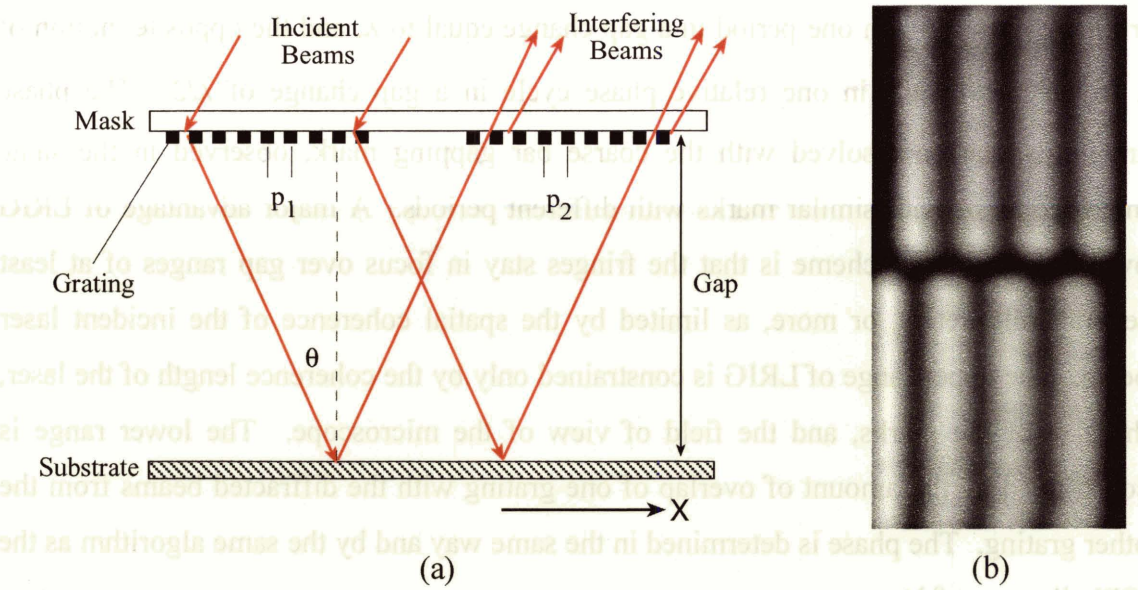


Fig. 4.9. (a) Diagram of long-range interferometric gap scheme. Interference of a reflected, forward-diffracted beam with a backdiffracted beam from a slightly different period results in a set of interference fringes. Two sets of marks, with p_1 and p_2 adjacent to each other, and a second set with the periods reversed, results in counterpropagating fringe sets that look identical to ISPI alignment fringes, but are dependent only on gap. (b) Interference fringes produced by two sets of gratings $\{p_1, p_2\}$ and $\{p_2, p_1\}$, with a 270 micron gap. One fringe cycle corresponds to a 320 nm change in gap. Phase ambiguity is resolved with geometric marks (as in Section 4.2), or by the spatial frequency content of the fringes, which depends on the amount of geometric overlap of the p_1 and p_2 gratings.

As we saw in the last few sections, a constant-period grating has both a backdiffracting component and a forward-diffracting component. The forward diffracting component reflects off the substrate back at the same angle as the backdiffracting beam, where they are both imaged by the ISPI microscope. Now consider a second grating with another period p_2 placed on the mask at some distance

from the first grating with period p_1. Both p_1 and p_2 are similar, e.g., $p_1 = 1.000$ and $p_2 = 1.025$ um, and diffract at or near the Littrow angle. Beams that forward-diffract from p_1, will reflect, and intercept the p_2 grating. At that position, the diffracted/reflected beams from p_1 overlap with the backdiffracted beams from the p_2 grating. Since $p_1 \approx p_2$, both beams are close enough in angle to return to the microscope, yet interfere to produce fringes. A second pair of gratings with the same periods, but in the opposite order, produces counterpropagating motion of the fringes as the gap changes. A single set of fringes cycles through one period in a gap change equal to λ, and the opposite motion of a fringe pair results in one relative phase cycle in a gap change of $\lambda/2$. The phase ambiguity can be resolved with the coarse bar gapping mark, observed in the same microscope, or with similar marks with different periods. A major advantage of LRIG over the geometric scheme is that the fringes stay in focus over gap ranges of at least several millimeters, or more, as limited by the spatial coherence of the incident laser beam. The upper range of LRIG is constrained only by the coherence length of the laser, the size of the marks, and the field of view of the microscope. The lower range is constrained by the amount of overlap of one grating with the diffracted beams from the other grating. The phase is determined in the same way and by the same algorithm as the ISPI alignment fringes.

Phase sensitivity implies fringe sensitivity that is $\lambda/2000$ of a fringe cycle, or 330nm/2000 = 0.2 nm. Phase ambiguity is avoided by geometric bar mark separation or simply by the amount of geometric overlap of the p_1 and p_2 gratings in the LRIG mark. Since the grating overlap determines the number of fringes that are visible, the latter method is preferred, and is amenable to frequency analysis, i.e., we encode coarse gap in the spatial frequency content of the image[24]. Such large-gap detection is useful in many applications outside of proximity lithography, such as scanning probe control, or three-dimensional stage positioning.

[24] Using the spatial frequency content of the fringe pattern has the non-trivial advantage that the coarse gapping method will not vary in focus over the range of gaps, as will occur with the geometric bar marks. Gap detection range is therefore not restricted to the focal depth of the microscope.

4.5 Summary

In this chapter we described three gapping schemes that are compatible with simultaneous alignment detection using ISPI microscopes. Large gaps are detected with either the geometric separation of diffracted and reflected beams resulting from a single small-area grating on the mask, or with the phase of a set of interference fringes that appear the same as ISPI alignment fringes, but are sensitive only to gap. Geometric gap detection yields <1 μm sensitivity, but the fringe-based scheme can detect <1 nm gap variations, by analyzing the spatial phase variations. The focal depth of the microscope does not restrict the range of this gapping method, since the interference fringes depend on the spatial coherence of the illumination.

Gaps less than 30 μm are detected with a scheme based on the phase between interferometric intensity oscillations in an inclined version of the Fabry-Perot interferometer. Absolute gap is obtained by comparison of intensity oscillations in multiple scans over the same range of gaps, using a different wavelength for each scan. Gap sensitivity is ~35 nm, which is sufficient for most lithographic purposes. However, this Chromatic Fabry-Perot scheme suffers from deficiencies related to scan repeatability and latency.

5 Transverse Chirp Gap Detection

Highly sensitive detection of small gaps is increasingly critical for nanopositioning and proximity or contact lithography. This chapter describes a novel interferometric method of measuring gaps between a mask and substrate, with <1 nm sensitivity, over a gap range between <1 μm and 30 μm. Gap is detected without scanning the mask height or the illumination spectrum. Instead, the method encodes gap in a unique combination of phase and frequency of a pair of interference fringes, produced from a pair of chirped checkerboards on the mask. The gapping mark is observed with the same microscope as the ISPI alignment marks, and at the same oblique angle.

5.1 Transverse Chirp Gapping (TCG)

The essential elements of a gap detection scheme, which we call Transverse Chirp Gapping (TCG), are illustrated in Fig. 5.1.

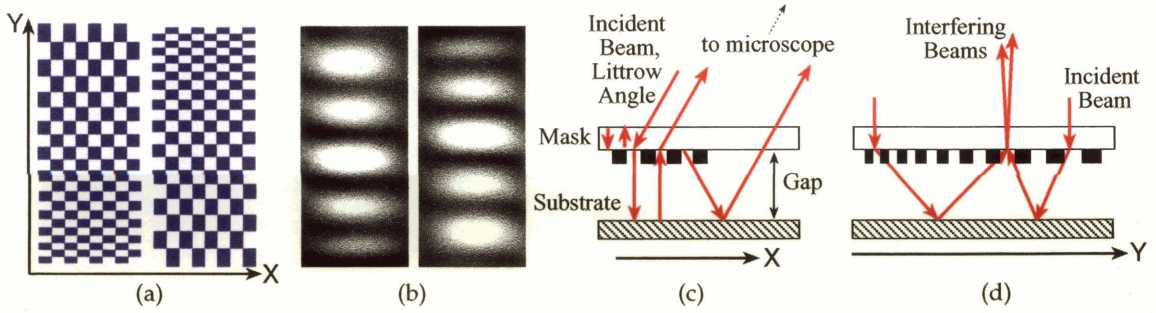


Fig. 5.1. Schematic of Transverse-Chirp-Gapping (TCG) gap-detection method. (a) Checkerboard mark with a chirp in the Y (transverse) direction and a constant period in the X (incident) direction. A second, adjacent chirped checkerboard produces fringes that move in opposite directions with changes in gap. (b) Image of interference fringes, as detected by an ISPI microscope. (c) Schematic in the incident plane illustrating the path by which light is returned to the microscope at the Littrow angle. (d) Schematic of diffracted beams in the transverse plane, showing an example of two incident beams that diffract, reflect, and rediffract, producing constructive interference. The distance between the two incident beams is called the interaction distance. The interaction distance can be viewed as an analog to the p_1, p_2 interaction of ISPI alignment fringes.

In TCG, a mark in the form of a two-dimensional grid, or checkerboard, is patterned on the mask, Fig. 5.1(a). Its spatial period along the X (incident) direction is uniform, and its spatial period along the Y (transverse) direction is non-uniform, i.e., chirped. The mark is illuminated with a single-wavelength beam at a Littrow angle of about 20° in the incident (X-Z) plane. The innovative feature in the TCG mark is that the chirped Y periodicity produces interference fringes from a single mark by interference of beams diffracted through a sweep of angles.

The chirped mark diffracts beams in a sweep of angles in the transverse plane. Subsequently, the beams reflect from the substrate, and rediffract from the mask at some distance from where they originally diffracted, Fig. 5.1(d). A portion of a rediffracting

beam returns in the incident plane towards an ISPI microscope, Fig. 5.1(c). In the transverse plane, interference of two or more such beams results in a fringe pattern, as shown in Fig. 5.1(b).

Points of constructive interference can occur only at periodic intervals along the length of the checkerboard, but shift in both displacement and separation when the gap is varied, resulting in simultaneous variation of fringe position and frequency. A sequence of TCG fringes at several gaps is shown in Fig. 5.2.

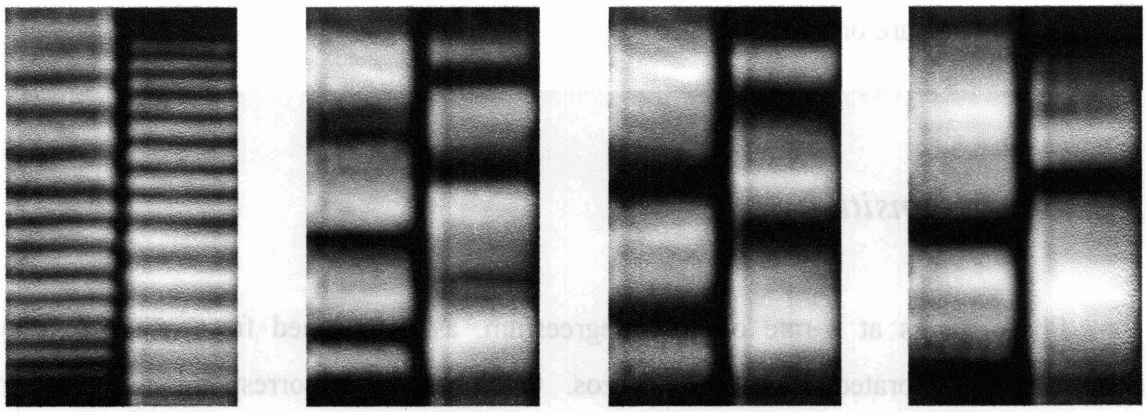


Fig. 5.2. Coarse gap detection demonstrated with TCG images at gaps of (a) 7 μm, (b) 2 μm, (c) 1.5 μm, and (d) 1 μm.

Fringe frequency is controlled by the so-called interaction distance along the chirp, i.e., the distance over which diffracted and reflected beams can interfere at a single point on the mask surface (Fig. 5.1(d)). The interaction distance increases with gap, causing more disparate spatial frequencies to interact, resulting in higher fringe frequencies. Conversely, as the gap is reduced, interfering beams traverse a smaller lateral distance, implying a smaller difference in interacting grating periods, and hence larger distances between regions of constructive interference. Experimentally, the rate of change in fringe

frequency is determined to be 1.62 fringes/μm, with $\lambda = 656$ nm and a chirp rate with the period given by

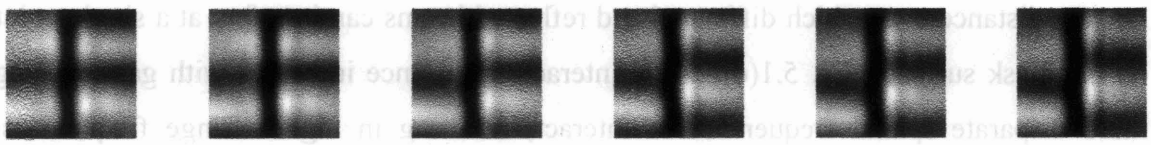
$$p(y) = (Ky + C)^{-0.25}, \tag{5.1}$$

where K and C are constants. Different fringe sensitivities as a function of gap are selected by design of the chirp rate.

Fringe formation does not require any marks on the substrate. TCG is therefore insensitive to misalignment, and can be used equally well for zero-level patterning, before any marks are on the substrate.

5.2 Phase Sensitivity

Phase varies at a rate of 1.79 degrees/nm, as determined from scans of gap, controlled by calibrated, closed-loop piezos. One phase cycle corresponds to a 200 nm gap change. Phase changes of ~53 degrees, corresponding to ~30 nm gap changes, are detectable by eye, and further sensitivity is achieved with the ISPI phase algorithm, capable of detecting <1/2000 of a fringe cycle. A minimum TCG phase increment corresponds to 200nm/2000 = 0.1 nm, indicating sub-1 nm gap sensitivity.



(a) (b) (c) (d) (e) (f)

Fig. 5.3. Fine gap detection demonstrated with portions of TCG images, starting at a 5 μm gap (a), and at gap increments of (b) +20, (c) +40, (d) +60, (e) +80, and (f) +100 nm, controlled by calibrated, closed-loop piezos. Using the ISPI frequency-domain algorithm, the phase disparity across the midline can be measured with a sensitivity corresponding to <1 nm gap increments.

Frequency detectivity of 1.62x0.2 = 0.324, or ~1/3rd of a fringe, is sufficient to overcome phase ambiguity.

5.3 TCG Simulation

A simulation was written to predict the behavior of the TCG fringes, as illustrated in Fig. 5.4.

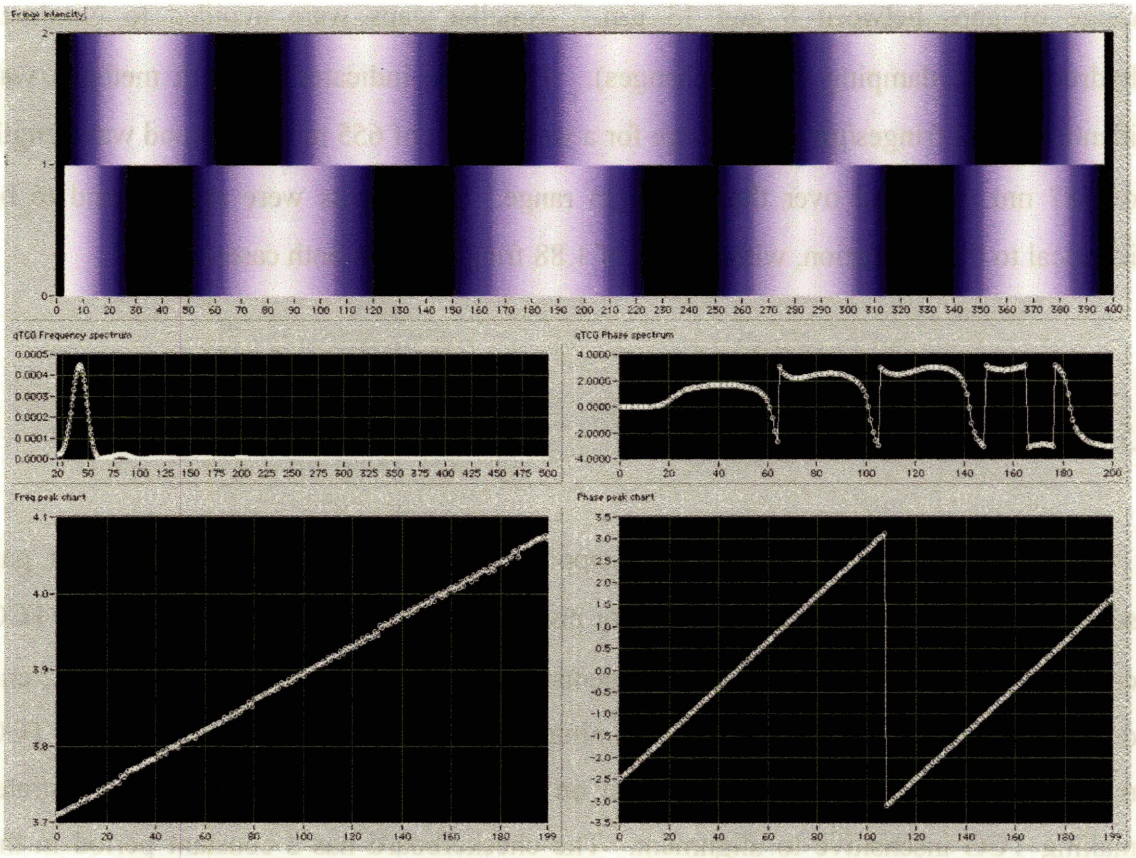


Fig. 5.4. TCG simulation front panel. (a) Simulated intensity plot from two sets of fringes. (b) Frequency and phase plots, indicating the narrow bandwidth of fringes produced from a quartic chirp rate. (c) Plots of intensity and phase over a range of gaps.

In the simulation, the grating is divided into 400 discrete segments (1 μm per segment). In each segment, the grating period is assumed to be a constant, where the local period is determined by the quartic chirp function, Eq. 5.1. The first-order diffraction angle is calculated from Eq. 3.3 and the diffracted beam traced along the path to the substrate, and on the return path to the mask. The phase contribution of each beam impinging on a particular segment on the mask is calculated from the optical path length. The intensity within each segment is calculated from the sum of phase contributions of multiple beams impinging on a given segment.

Results of the simulation were compared with TCG frequency measurements over a range of gaps between 8 and 13 μm. (Smaller gaps were avoided to eliminate hydrodynamic damping of gap changes). The slope indicated by both methods was identical: 1.62 fringes/μm gap change for a wavelength of 655 nm. A second wavelength of 687 nm was used over the same gap range. The results were again found to be identical to the simulation, with a slope of 1.88 fringes/μm in both cases.

5.4 Summary

In this chapter we described a gapping scheme that demonstrates <1 nm gap sensitivity at small gaps typically used for proximity lithography (1 μm to 30 μm). TCG consists of direct, site-by-site measurements, made using the same optics as ISPI, and does not require scanning of gap or illumination spectrum for gap detection. The TCG marks are composed of a checkerboard on the mask, and no mark on the substrate, making TCG insensitive to alignment. The checkerboard has a constant period in the incident plane, and a chirped period in the transverse plane. Two such marks with their chirps in opposite directions result in fringes that counterpropagate when the gap is changed. Their relative phase encodes position, as in ISPI. A distinctive feature of the TCG fringes is that their spatial frequency varies simultaneously with spatial phase,

allowing an unambiguous encoding of gap in a unique combination of frequency and phase.

TCG gap and ISPI alignment fringes are typically detected in the same image, permitting simultaneous measurement of Z and either X or Y. Three ISPI microscopes measure TCG and ISPI at three positions, yielding simultaneous six-axis position detection: three Z measurements (and derivation of mask tip and tilt), one X and two Y measurements (and derivation of in-plane rotation).

TCG can be used in conjunction with Long-Range Interferometric Gapping (LRIG), described in Section 4.4, to achieve sub-nanometer detectivity over a range from <1 μm to >500 μm. Neither TCG nor LRIG require scanning to detect gap, making them ideal for feedback control of gap.

6 The Chirped Talbot Effect

This chapter introduces a new variant of the Talbot effect that utilizes near-field period doubling from a chirped period grating to produce non-constant-period fringes in resist. The phase and frequency of the fringes in resist are matched to finite-difference time domain (FDTD) simulations to determine the absolute gap at the time of exposure. The sensitivity of this Chirped Talbot Effect is <10 nm, an improvement of more than two orders of magnitude over the ordinary Talbot effect. The Chirped Talbot Effect can be used in conjunction with Transverse Chirp Gapping. The two forms of gap measurement are complementary since one is suited to measuring and setting gap prior to exposure, and the other is ideal for confirmation of the gap that existed during exposure.

6.1 *The Talbot Effect*

When a periodic structure, such as a grating, is illuminated with spatially and temporally coherent light, constructive interference results in intensity patterns that mimic the original period at certain distances beyond the grating, Fig. 6.1. This is known as self-imaging, or the Talbot effect [6.1].

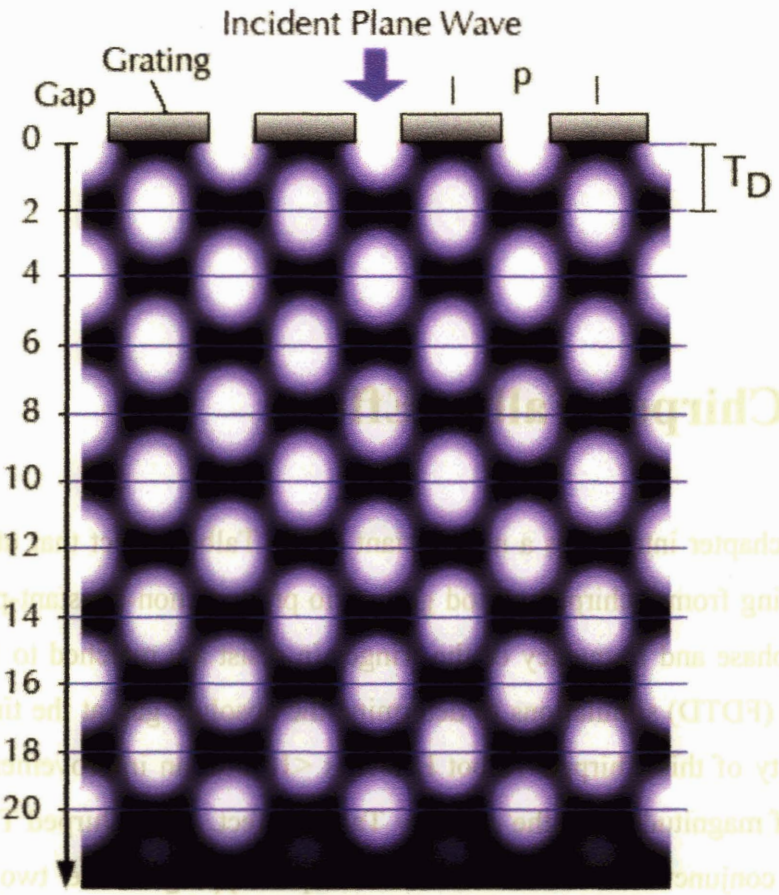


Fig. 6.1. Two-dimensional FDTD simulation of the Talbot effect with a grating of period p = 1.0 μm, λ = 0.5 μm. The Talbot distance $T_D = p^2/\lambda$ refers to the minimum distance between replications of the fundamental spatial frequency. Note that the first Talbot distance occurs at 2.0 μm. Even multiples of the Talbot distance reproduce the original spatial phase and frequency, while odd multiples reproduce the spatial frequency (i.e., the fundamental) with a π-phase shift. Doubling of the fundamental frequency occurs at odd integer multiples of $T_D/2$. Surfaces of constant spatial frequency are parallel to the grating.

The minimum distance between replications of the fundamental spatial frequency is the Talbot distance,

$$T_D = \frac{p^2}{\lambda} \tag{6.1}$$

where p is the grating period and λ is the wavelength of illumination. For instance, with p = 1.0 μm, and λ = 0.5 μm, T_D = 2.0 μm. In between these Talbot distances are gaps with double the fundamental spatial frequency. Since the TCG measurement is capable of detectivity of the order of 1 nm, the standard Talbot effect is inadequate for verification of TCG fringe interpretation.

6.2 Chirped-Talbot-Effect Gap Verification

In the standard Talbot effect, the distance T_D is a constant for a given grating period, p. In the Chirped-Talbot effect (CTE), with the grating period varying along the length of the mark, a variety of spatial frequencies are present at a given gap. Figure 6.2 depicts such a chirped grating, and the associated intensity pattern. The inclined planes at integer multiples of the Talbot distance, T_D, are referred to as Talbot surfaces. In Fig. 6.2 the Talbot surfaces appear as straight, slanted lines due to the narrow region of the grating that is shown and the small chirp rate.

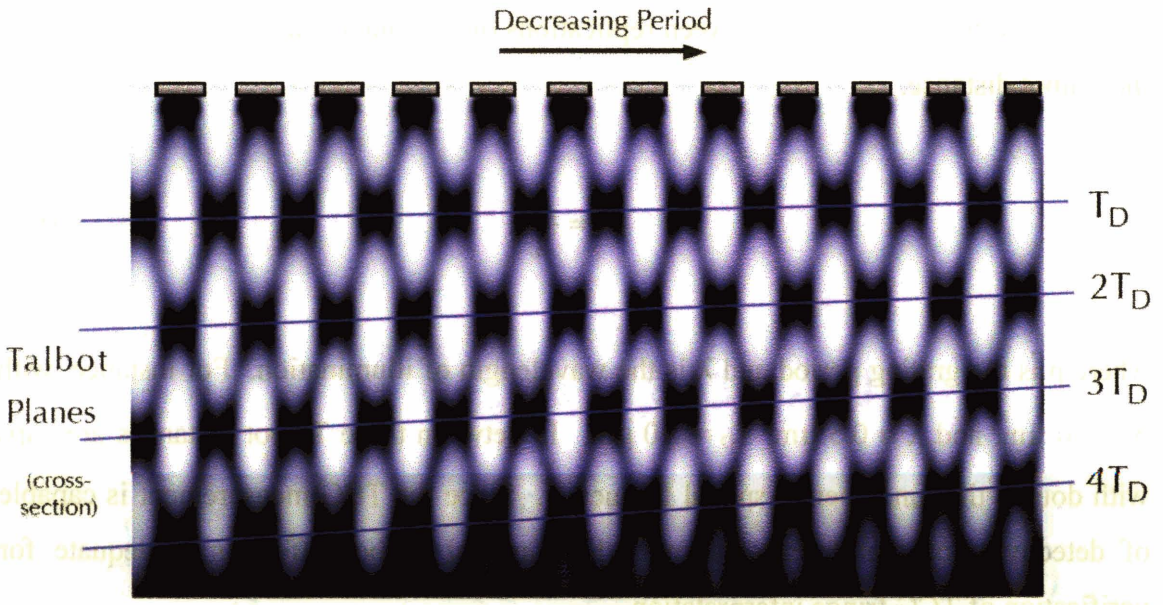


Fig. 6.2. FDTD simulation of the Chirped-Talbot effect. A variation in grating period results in an inclination of the Talbot surfaces. With a chirped grating a number of Talbot surfaces can intersect a substrate plane that is parallel to the grating plane.

By increasing the chirp rate and the length of the mark, multiple Talbot surfaces can intersect a substrate plane that is parallel to the mask or template. Figure 6.3(a) shows the standard Talbot effect, Fig. 6.3(b) shows a cross-section through Talbot surfaces for a linear chirp, and Fig. 6.3(c) shows a cross-section through Talbot surfaces for a quartic chirp.

In the latter two cases, the spacing between Talbot surfaces is a variable that one can adjust via the chirp function. Wherever a Talbot surface intersects the substrate, the grating fundamental is reproduced in the resist, and one can observe a checkerboard pattern identical to the TCG checkerboard mark, i.e., a two-dimensional Talbot effect. In between the locations where the Talbot surfaces intersect the substrate plane, frequency doubling occurs. At these locations we typically observe complete removal of the photoresist. As one goes from one of the Talbot-surface intersections to a location of frequency doubling at an intersection with a half-Talbot surface, the size of the

checkerboard dots shrinks, producing a continuous grayscale effect when viewed in a low-NA normal-incidence optical microscope.

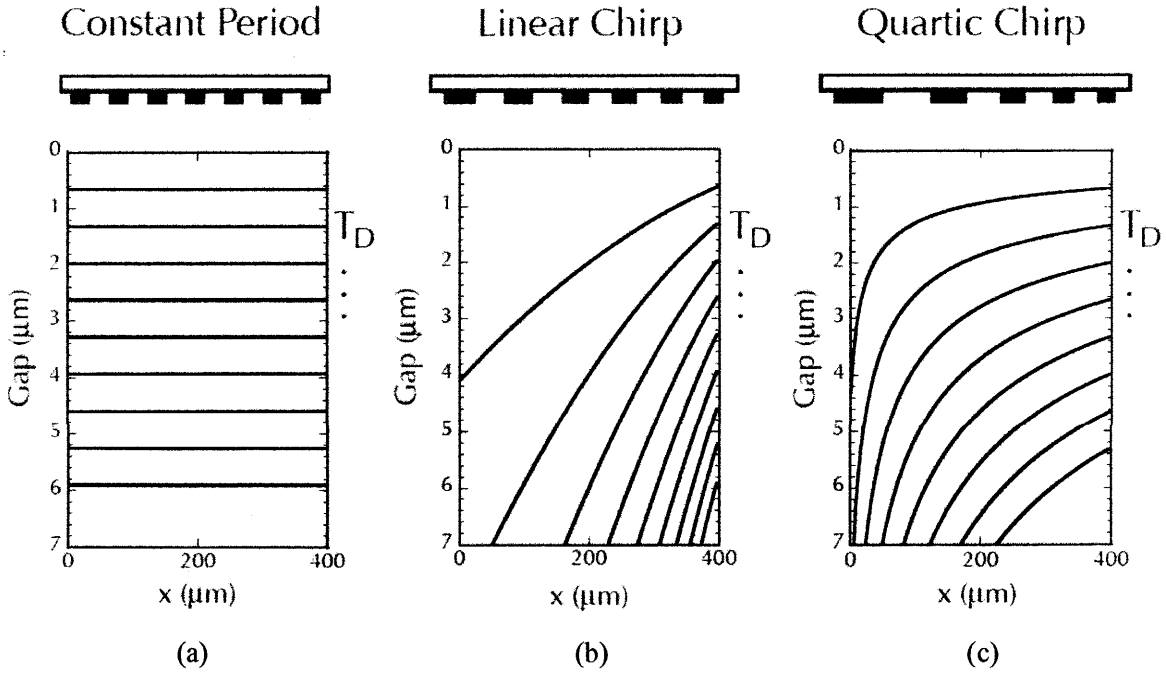


Fig. 6.3. Talbot surfaces from (a) a fixed-period grating, p = 0.8 μm, (b) a linear-chirped grating, p(y) = Ky + C, and (c) a quartic-chirped grating, $p(y) = (Ky + C)^{-0.25}$, assuming λ = 488 nm. Note that each curve corresponds to a gap increment of $T_D/2$. Compared to the linear-chirped grating, the quartic-chirped grating provides a narrower bandwidth of spatial frequencies at a given gap, which simplifies analysis. Plots are based on a model that assumes a locally-constant grating period.

6.3 Chirped-Talbot Simulations

The curves in Fig. 6.3 are approximations that assume a locally-constant grating period. Deviations from this approximation become evident as the gap and chirp rate are increased. A rigorous calculation of the intensity as a function of gap was carried out using a finite-difference time domain (FDTD) electromagnetic simulation, Fig. 6.4. The FDTD simulation was performed on a 2D grid of points at 20 nm intervals over a 400 μm-long mark (400μm/0.02μm=20,000 points along each row in the X direction), and up to a 7 μm gap. The FDTD was performed in TEMPEST v. 6.1, using a custom Matlab

front end. The disparity between the FDTD simulation and the local-interaction approximation is illustrated in Fig. 6.4, with the lines from Fig. 6.3(c) superimposed.

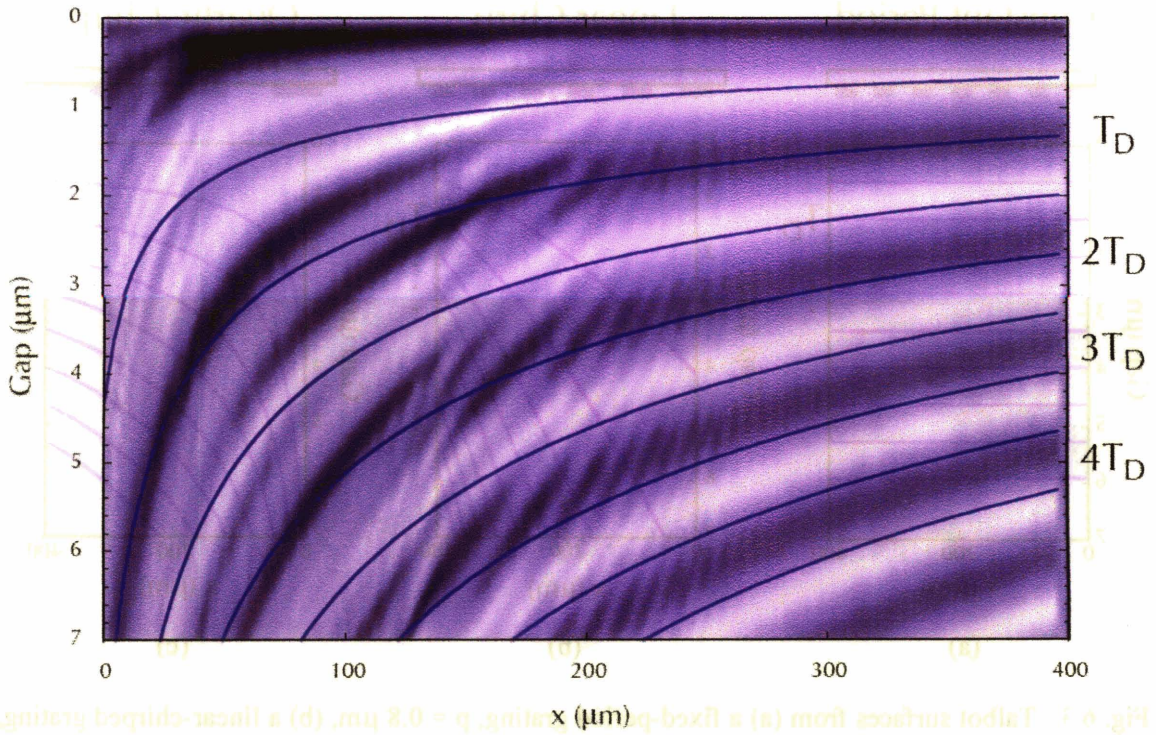


Fig. 6.4. Two-dimensional FDTD simulation of intensity from a quartic-chirped grating, using the same constants and wavelength as in Fig. 6.3(c). Superimposed lines from Fig. 6.3(c) indicate the local-interaction approximation to Talbot surfaces. Deviations from the local approximation are evident for increasing gap and increasing chirp rate. Since there are 20,000 points in each row of the simulation, the plot above shows a downsampled version. Regions of doubled spatial frequency are shown as high intensity, and regions of low spatial frequency as low intensity, in effect illustrating the spatial frequencies that are expected in resist as a moiré of the intensity at each point.

6.4 Chirped-Talbot Exposures

Using the Chirped-Talbot effect (CTE), resist at several gaps was exposed (in a configuration as in Fig. 6.5) with normally-incident argon-ion laser light (interference-filtered to a single line at $\lambda = 488$ nm).

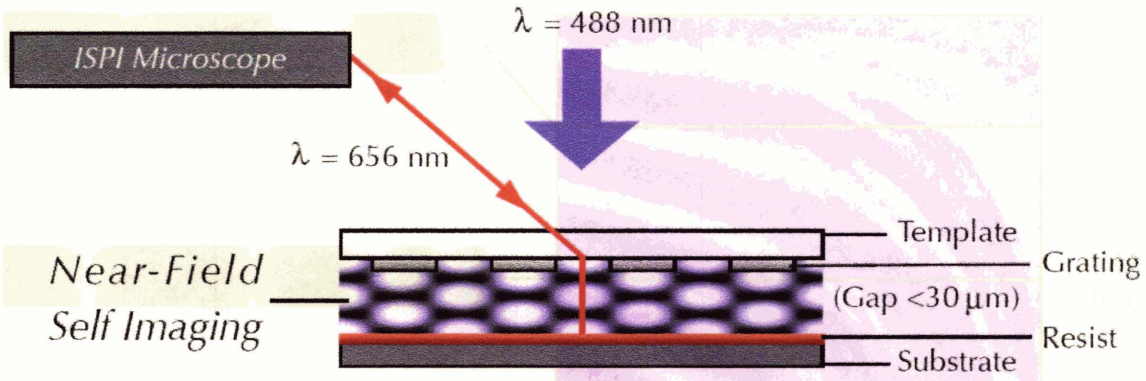


Fig. 6.5. Schematic of an experiment with simultaneous observation of TCG fringes and exposure in resist with a 488 nm Ar^+ laser. The near-field self-imaging intensity nodes are recorded in resist and a halftone effect due to varying dot size in the resist yields a smooth grayscale distribution when viewed with a low-NA, normal-incidence optical microscope.

After exposure and development, optical micrographs were taken of the fringes in resist. Absolute gap was determined by the spatial frequency and phase content of the exposed regions. Figure 6.6 shows fringes in resist resulting from exposures at gaps of 2.0, 4.9, and 7.4 μm. Since we expose two mirror-image TCG checkerboard marks (i.e., oppositely-directed chirps) at the same time, we achieve enhanced sensitivity. Figure 6.7 shows an example of phase analysis. Absolute gap with sub-10 nm sensitivity is found by matching the frequency spectra of the exposed features to FDTD simulations and measuring the phase disparity between resist fringes. The phase sensitivity of the CTE fringes is ~0.02 degrees/nm, which is adequate to verify the interpretation of the TCG fringes to better than 10 nm.

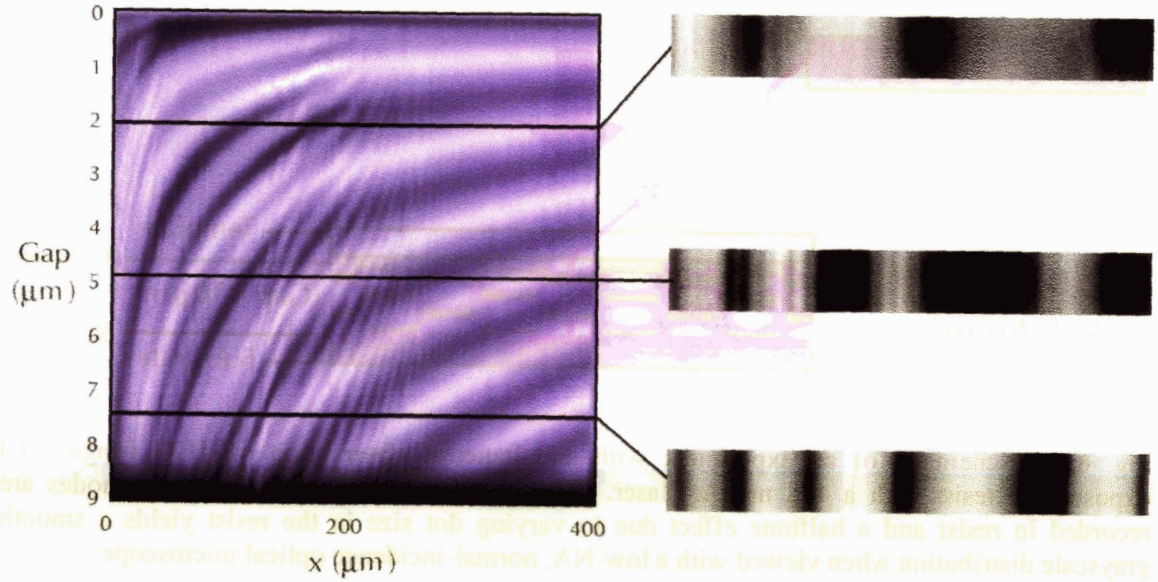


Fig. 6.6. Chirped-Talbot Effect FDTD simulation (a), and exposed resist at several gaps. Fringes in resist for gaps of (b) 2.0 μm, (c) 4.9 μm, and (d) 7.4 μm.

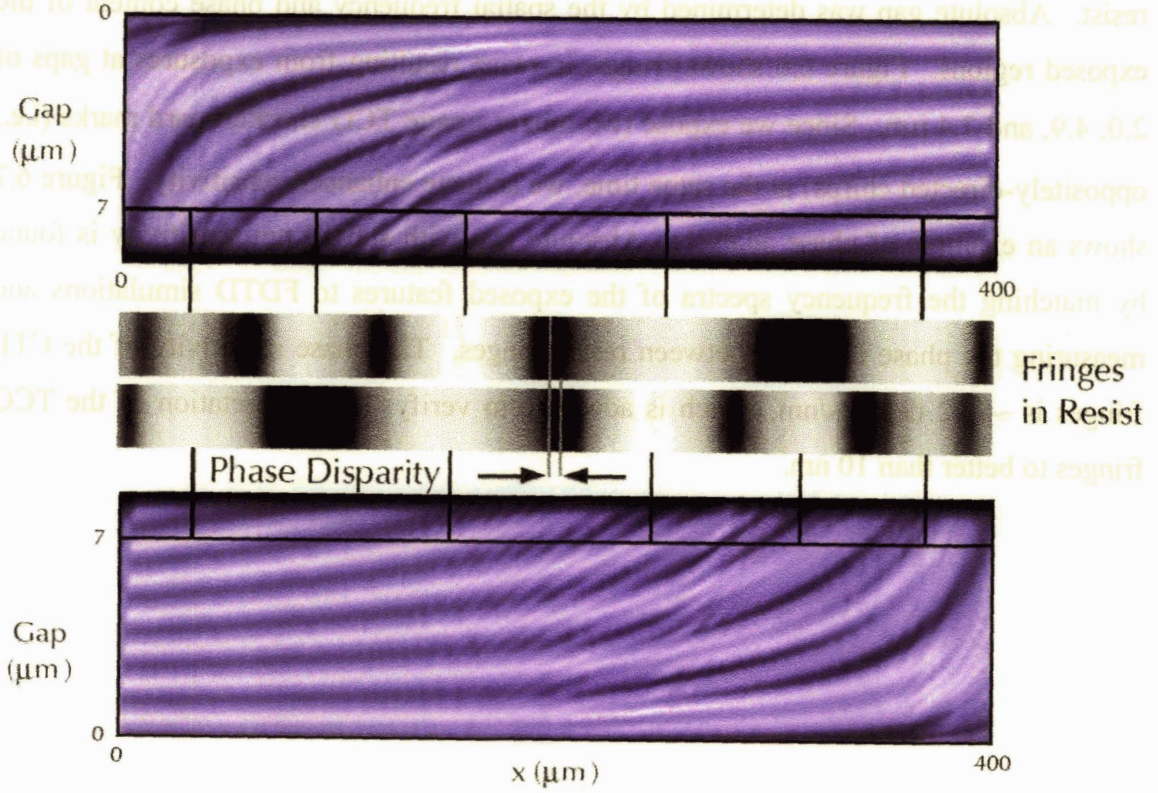


Fig. 6.7. Diagram of the Chirped-Talbot Effect FDTD for both sides of a chirped-checkerboard mark, and the associated phase disparity between the fringes exposed in resist.

6.5 *Summary*

We have described a means of measuring and verifying the gap between a mask and a substrate in proximity using a single mark on the mask. The novel aspect of this method is the combination of two independent means of ascertaining gap in a single mark: in one form, the spatial phase and frequency of interference fringes are observed with an oblique-incidence microscope using a non-exposing wavelength (TCG), and, in the other form, the spatial phase and frequency of fringes from the same mark are observed *ex post facto* in exposed resist (CTE). Gap is encoded in spatial phase and frequency in an identical manner in both forms. The forms are complementary since one is suited to measuring and setting gap prior to exposure, and the other is ideal for confirmation of the gap that existed during exposure. The Chirped Talbot fringes arise from surfaces of constant spatial frequency in an interference pattern (Talbot surfaces). The pattern of fringes is controlled via the chirp rate of the mark on the mask. This gapping method is fully compatible with the analogous Interferometric-Spatial-Phase Imaging (ISPI) alignment method, and meets the requirements of a plethora of applications including template or mask leveling, mask control at small exposure gaps, or scanning tip control, among others.

7 Effects of Perturbations

This chapter examines the effects of numerous factors that might influence or degrade the accuracy of ISPI measurements. Overlayers of resist, polysilicon, and aluminum are tested and found to affect the fringe contrast, but not the spatial phase. The effects of grating defects are evaluated using purposeful imperfections in ISPI alignment marks. ISPI phase is measured as a function of microscope rotations in the incident and transverse planes. Specialized marks are introduced to verify the viewing angles, using interferometric effects. A scheme for automatic correction of small axial camera rotation errors is implemented. Readings from two separate ISPI microscopes, looking at the same alignment mark, were compared and found to be consistent, with mean differences <1 nm.

7.1 Process Layers

Numerous alignment schemes described in Chapter 2 promised nanometer-level detection, yet in practice fell short of expectations, due in large part to the effects of process coatings. When alignment marks on the substrate are covered with various grown and deposited layers, such as resist, polysilicon, and aluminum, the phase and amplitude of the reflected light are changed due to scattering and interference of beams from the multiple interfaces. This usually alters the alignment signal, leading to overlay errors.

We expect a spatial-phase detection scheme such as ISPI to be immune to many of these potentially disturbing influences. In an experiment, an ISPI alignment mark was etched into silicon, and half of the mark was covered with an overlayer (Fig. 7.1). The mark was observed with an ISPI microscope, and the resulting fringe patterns for the upper and lower halves of the mark were analyzed separately, using the ISPI phase algorithm.

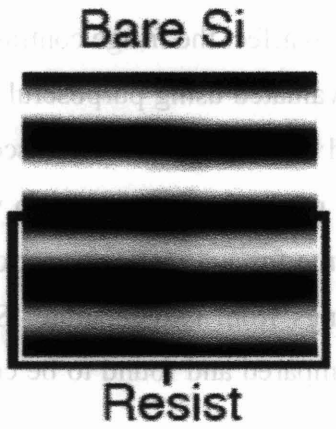


Fig. 7.1. ISPI image of fringes from an alignment mark etched into a substrate. Half of the mark was covered with an overlayer. The other half was left bare. Experiments were done in this manner with overlayers of resist, polysilicon, and aluminum.

The effects on alignment of three types of overlayers—photoresist, polysilicon, and aluminum—were investigated. Substrate alignment marks consisted of gratings p_1 = 3.7 μm and p_2 = 4.0 μm, etched 100 nm deep into a Si substrate. The marks, which occupied an area 200x200 μm, were covered with a given overlayer, and then that overlayer was removed from one half of the substrate alignment mark (by exposure of coarse features aligned to the mark, and subsequent etching of the polysilicon and aluminum, or for resist, simple development and removal). Resist and polysilicon decreased fringe contrast, while aluminum increased contrast. The mask-to-substrate gap was typically 10 μm although any setting from 1–50 μm would also have been appropriate.

Figure 7.2 displays the results of measurements of the difference in the spatial phase discontinuity in the half of the alignment mark with the overlayer and the half without the overlayer.

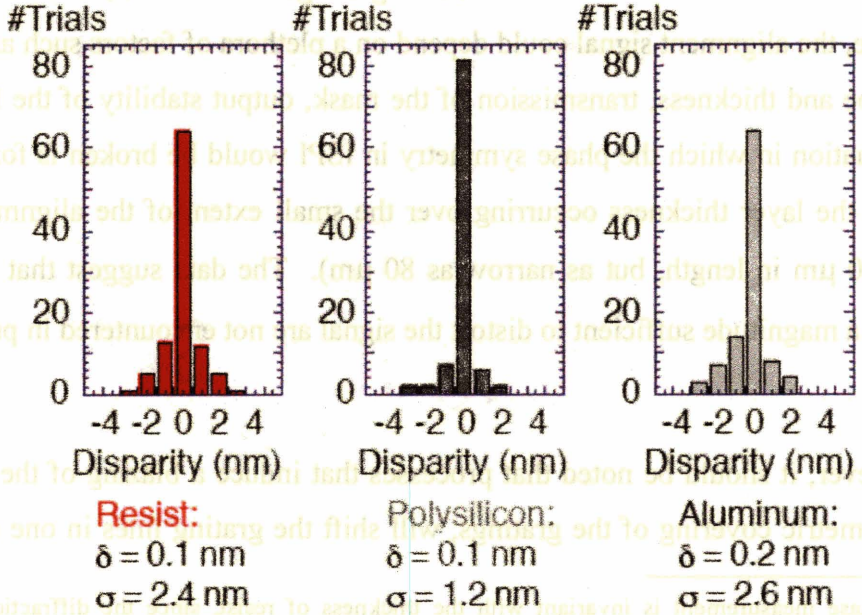


Fig. 7.2. Immunity of ISPI to overlayers of resist, polysilicon and aluminum. Layers of resist, polysilicon, and aluminum are shown to cause minimal degradation of the alignment signal. Mean alignment errors are less than the detectivity of the alignment measurements.

The 100 measurements shown in Fig. 7.2 for each overlayer yielded a mean offset, δ, as well as a standard deviation, σ. The measured effects of the overlayers were $\delta = 0.1$ nm, $\sigma = 2.4$ nm for 1 μm of photoresist[25], $\delta = 0.1$ nm, $\sigma = 1.2$ nm for 300 nm of deposited polysilicon, and $\delta = 0.2$ nm, $\sigma = 2.6$ nm for 20 nm of deposited aluminum.

The significance of the results is that an overlayer *does not* appear to affect the *spatial phase* of the interference pattern, although it does affect the image contrast. Since we signify alignment by the matching of the spatial phase across the midline of the CCD image, the effect of overlayers on alignment is negligible; in fact, the mean error is less than the detectivity of the measurement.

We have demonstrated that any overlayer which causes a phase change on one side of the mark, will also, in all likelihood, affect the phase on the other side of the mark in exactly the same way since each set of fringes is formed by the interference of symmetric diffractions from mask and substrate gratings. In contrast, in an amplitude-based scheme, the alignment signal could depend on a plethora of factors such as the gap, overlayer type and thickness, transmission of the mask, output stability of the laser, etc. The only situation in which the phase symmetry in ISPI would be broken is for extreme variations in the layer thickness occurring over the small extent of the alignment mark (typically 400 μm in length, but as narrow as 80 μm). The data suggest that thickness variations of a magnitude sufficient to distort the signal are not encountered in practice.

However, it should be noted that processes that induce a blazing of the gratings, i.e., an asymmetric covering of the gratings, will shift the grating lines in one direction.

[25] The ISPI phase measurement is invariant with the thickness of resist, since the diffraction angle is independent of the layer thickness. Since the two halves of the ISPI mark are in close proximity (typically ~200 μm) the only variation of some concern to ISPI would be a ramp in the thickness of resist across the mark. Resist is typically flat and consistent in thickness to the nanometer level over the dimensions of the mark, and this is not observed to be a problem in practice.

This is equivalent to moving the substrate gratings, which we know will cause the observed fringes to shift in opposite directions. Thus, in the special case of asymmetric overlayers, we expect a spatial offset. However, the ability of ISPI to detect arbitrary offsets with equal accuracy should allow one to correct for a systematic error caused by an asymmetric covering of the grating.

7.2 Grating Defects

In any kind of moiré alignment scheme, one would expect the alignment signal to be less sensitive to defects than a single line in a conventional amplitude alignment scheme. This is clear from statistics – the accuracy of the estimator increases with the size of the sample set. In ISPI the sample number is large: for example, marks contain up to 400 lines on the mask, and 400x200 = 80,000 boxes on the substrate checkerboard, producing one pair of moiré fringes. One would expect deviation in position of one or a few grating lines, or other defects, to have a negligible effect on the phase of the fringe pattern.

Since ISPI derives many of its strengths from the design of the diffraction structures, it is reasonable to examine how imperfections in the structures affect the quality of the alignment signal. We do this with purposeful imperfections written in the gratings: line omissions, line shifts, grating region shifts, and line-space ratio variations.

In the mark designed for this experiment (Fig. 7.3), there are four grating sections on each side of a central reference grating. The substrate marks consist of a conventional checkerboard, with 1 μm period in the incident plane, and 2.05 μm period in the transverse direction. The reference region of the mark has a 2.05 μm-period grating on the mask and a 2.00 μm period checkerboard on the substrate. Mask marks contain eight test grating strips (each 45x400 μm), starting with a defect-free grating strip immediately

to the right of the reference grating. Increasing amounts of the defect occur in gratings further to the right, and wrapping around to the opposite side. This arrangement makes the maximum and minimum amount of defects occur adjacent to the center reference grating. The nominal period in each grating strip is 2.00 μm. Figure 7.3 shows the design of one such mark.

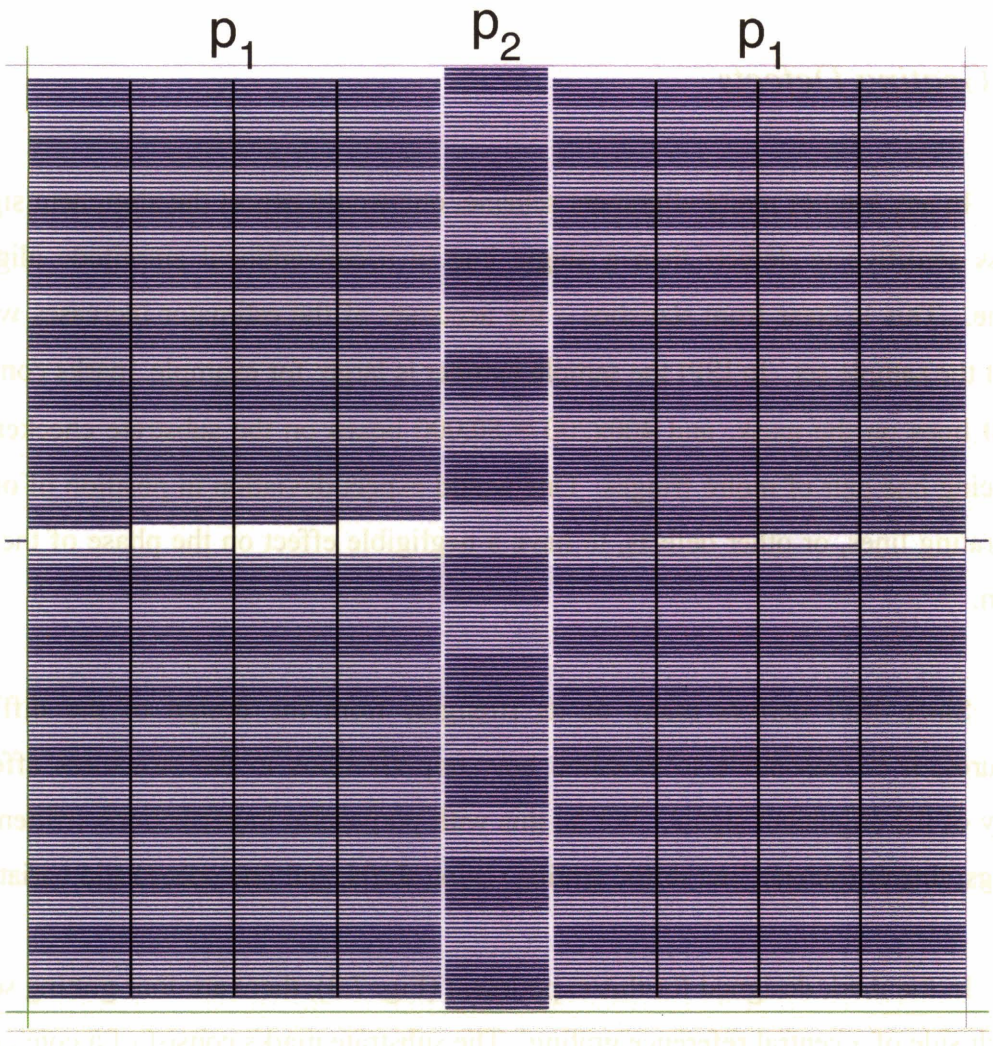
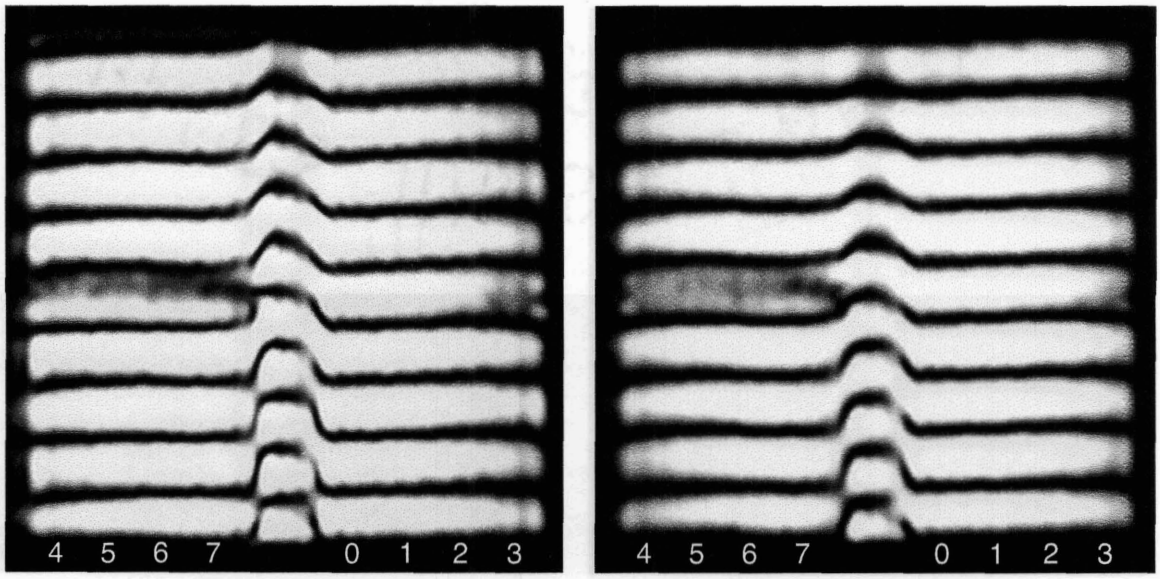


Fig. 7.3. NanoWriter KIC design of a defect test mark in the NanoWriter layout program. Eight grating regions are utilized, each with one more missing grating line. A grating region is 45x400 μm. A central grating serves as a phase reference. The overall width of the mark is 409.6 μm. (The moiré effect observed in this figure is an artifact of the resolution in the display screen.)

7.2.1 Grating Line Omissions

Using the mark shown in Fig. 7.3, and a checkerboard mark on the substrate, we observed a set of fringes, shown in Fig. 7.4.



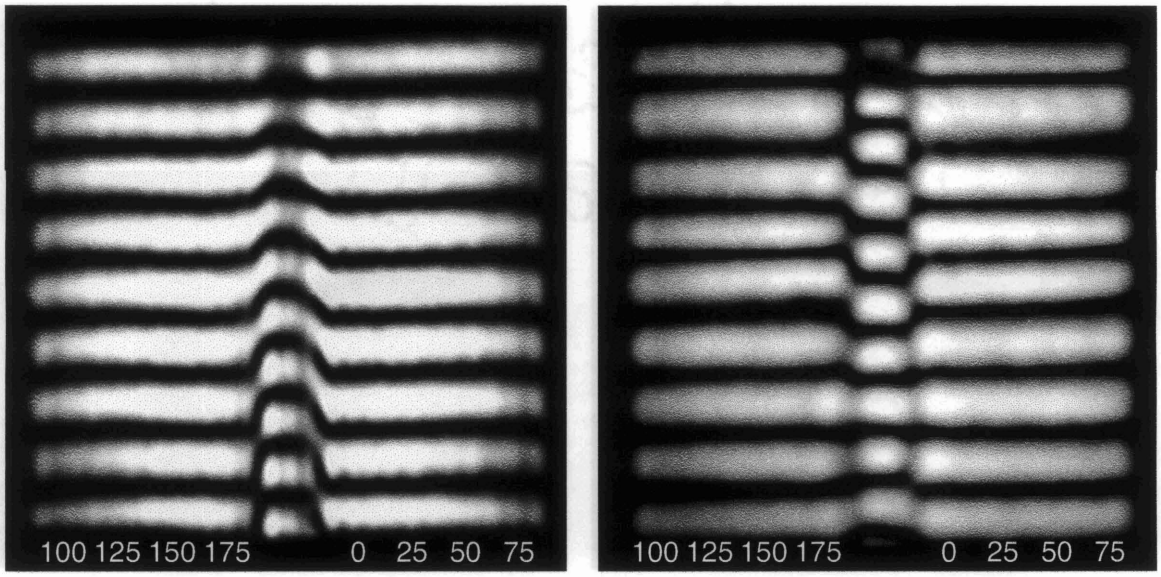
(a) (b)

Fig. 7.4. Images of gratings with omitted grating lines on the mask. The number of omitted lines along the horizontal increase from 0 to 3 on the right, and from 4 to 7 on the left. (a) and (b) are images of the same mark, with different fringe positions, indicating how the defects affect the fringes as they move. Grating periods are $p_1 = 2.00$ μm and $p_2 = 2.05$ μm.

The omitted grating lines are evinced as dark, stationary regions. As the fringes move during alignment, the grating omissions constitute a stationary dip in intensity. The effect of the omission may be unobservable if it coincides with an intensity minimum of the undistorted fringe pattern, or may reduce contrast in one fringe in the pattern, as shown in Fig. 7.4(b). It is interesting to note that omission of one or two lines, amounting to up to 10% of the 41 μm fringe period, is not discernable.

7.2.2 *Grating Line Shifts*

Shifts of a single grating line and a group of lines at the midpoint of the mark by successive e-beam step increments of 25 nm are shown in Fig. 7.5.



(a) (b)

Fig. 7.5. (a) Image of fringes from a test mark with a single grating line in the middle of the mark shifted in the +Y direction by increasing amounts of 25, 50, 75 nm on the right, and 100, 125, 150, and 175 nm on the left. (b) Image of a test mark with a group of grating lines shifted by the same amounts. One group is composed of 20 grating lines, equivalent to a 41 μm fringe period.

In Fig. 7.5(a), the effect of a shift of a single grating line is effectively unobservable, even in shifts as large as 175 nm.

Moving a section of the grating, as shown in Fig. 7.5(b), causes a prominent shift of one fringe by an amount that appears to be in proportion to the grating shift.

7.2.3 *Line-Space Ratio Variations*

Figure 7.6 displays fringes resulting from variation in the line-space ratio on the mask gratings.

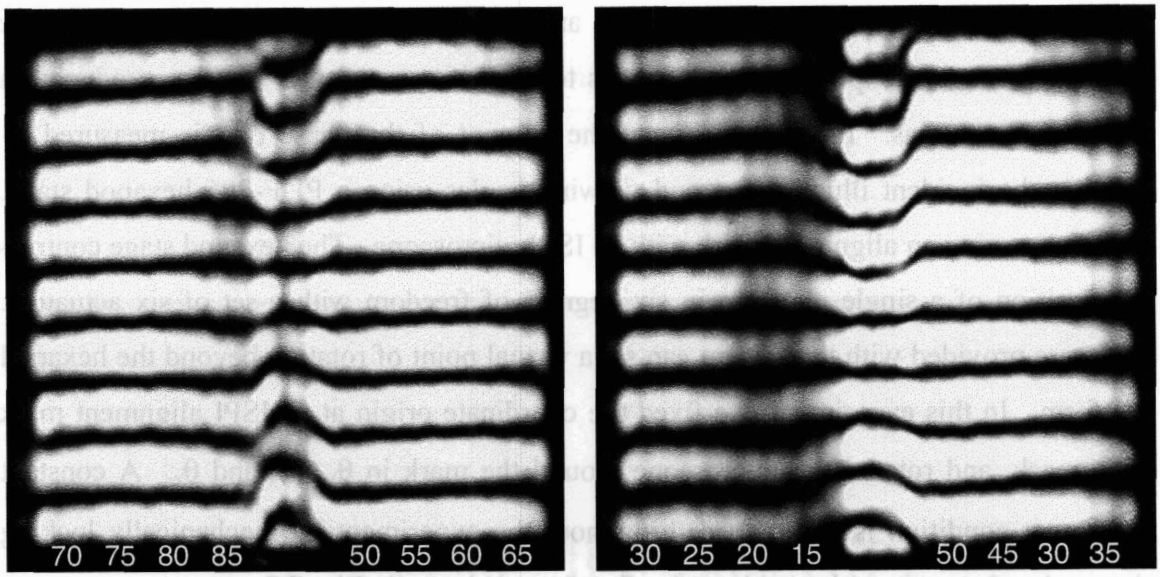


Fig. 7.6. Images of fringes from a line-space ratio test mark. (a) Fringes resulting from increasing duty cycle: 0.50, 0.55, 0.60, 0.65, 0.70, 0.75, 0.80, and 0.85. (b) Fringes resulting from decreasing duty cycle: 0.50, 0.45, 0.40, 0.35, 0.30, 0.25, 0.20, and 0.15. Numbers at the bottom of the figures indicate the percentage of the period occupied by an opaque line.

In Fig. 7.6 (a) the opaque linewidth is increasing, from 50% of the period, up to 85%. The fringes appear to have essentially constant intensity throughout the range of linewidths, even with 85% linewidth, in which case the clear aperture between lines is only 0.3 μm, less than half of the 690 nm illumination wavelength. The diffraction efficiency appears to be more adversely affected by decreasing line-space ratios, as in Fig. 7.6(b). Overall, the images suggest that a very wide latitude is possible in the line-space ratio, indicating robustness to process variations.

7.3 Oblique Viewing Angles

In Chapter 3, we alluded to the requirements of symmetric diffraction in ISPI fringe formation. Littrow-angle imaging breaks symmetry in one direction, but symmetry is required within the transverse plane to maintain balance between the two sets of ISPI fringes. Since the diffraction beam paths are normally symmetric on each side of the incident plane, we expect the ISPI fringes to vary if symmetry is broken, producing a constant phase bias. In an experiment, the amount of the phase bias is measured by rotating the incident illumination and viewing angles using a PI F-206 hexapod stage, while observing an alignment mark with an ISPI microscope. The hexapod stage controls the position of a single platform in six degrees of freedom with a set of six actuators. Software provided with the system can set a virtual point of rotation beyond the hexapod platform. In this experiment, we fixed the coordinate origin at an ISPI alignment mark on a mask, and rotated the microscope around the mark in θ_x, θ_y, and θ_z. A constant alignment condition is maintained throughout the experiment by mechanically locking the mask-substrate position with studs. Results are shown in Fig. 7.7.

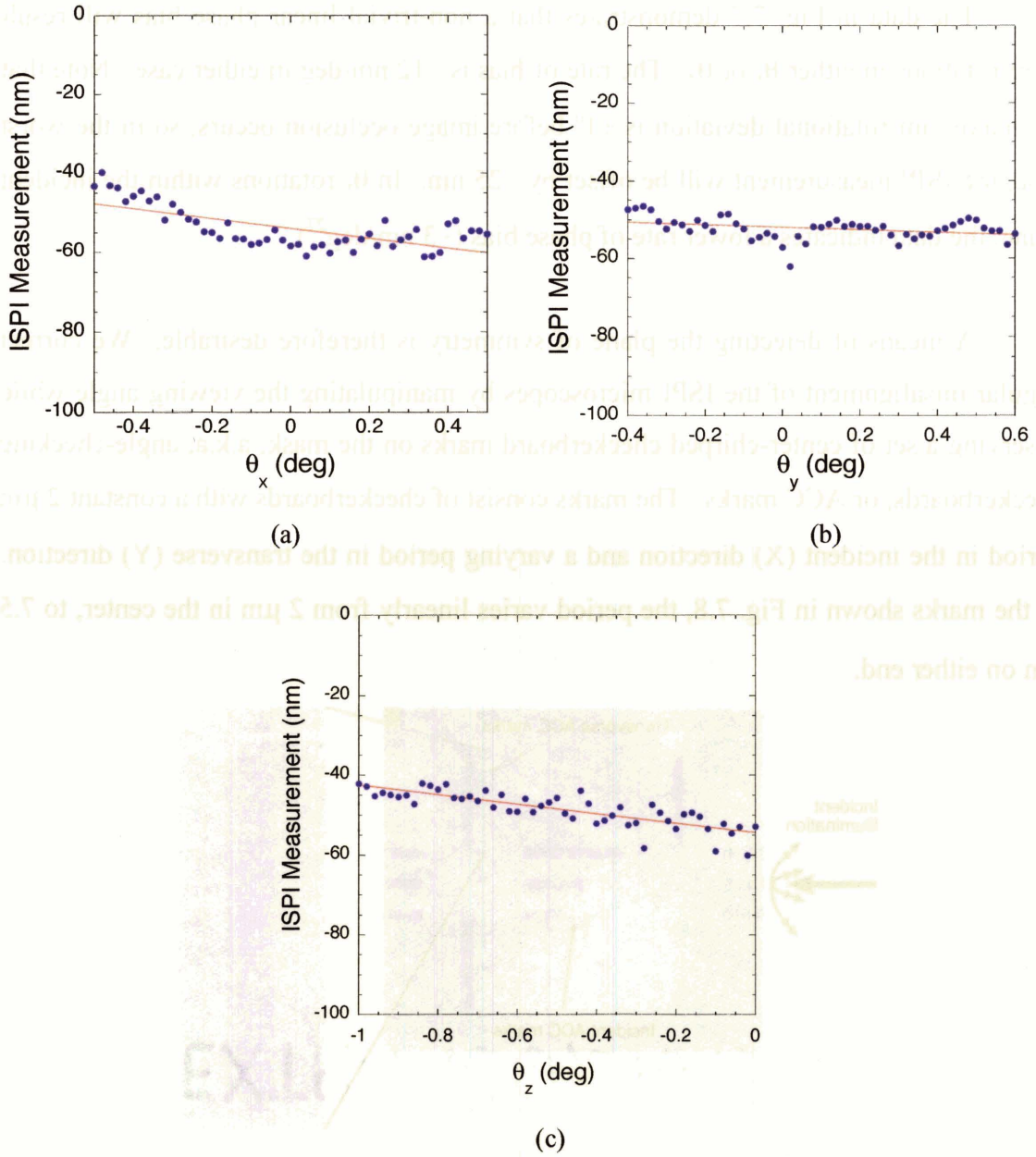


Fig. 7.7. Plots of ISPI measurements taken over a 1° range of angles by adjusting the angles of an ISPI microscope with a PI F-206 hexapod stage[26]. The incident plane is defined by the X and Z axes. (a) ISPI measurements during a rotation of θ_x. The slope of the linear fit is -12.4 nm/deg. (b) ISPI measurements during a rotation of θ_y. The slope of the linear fit is -3.3 nm/deg. (c) ISPI measurements during a rotation of θ_z. The slope of the linear fit is -11.9 nm/deg. The data suggest angular coupling between incident and transverse angles in the stage, as well as some jitter in the stage. The nominal minimum angular step size of the stage is 0.01°.

[26] Many thanks to Jim Gareau and PI for the extended loan of the hexapod.

The data in Fig. 7.7 demonstrates that a non-trivial linear phase bias will result from rotations in either θ_x or θ_z. The rate of bias is ~12 nm/deg in either case. Note that the maximum rotational deviation is ±1° before image occlusion occurs, so in the worst case, the ISPI measurement will be offset by ~25 nm. In θ_y rotations within the incident plane, the data indicates a lower rate of phase bias (~3 nm/deg[27]).

A means of detecting the plane of symmetry is therefore desirable. We correct angular misalignment of the ISPI microscopes by manipulating the viewing angle while observing a set of center-chirped checkerboard marks on the mask, a.k.a, angle-checking checkerboards, or ACC marks. The marks consist of checkerboards with a constant 2 μm period in the incident (X) direction and a varying period in the transverse (Y) direction. In the marks shown in Fig. 7.8, the period varies linearly from 2 μm in the center, to 7.5 μm on either end.

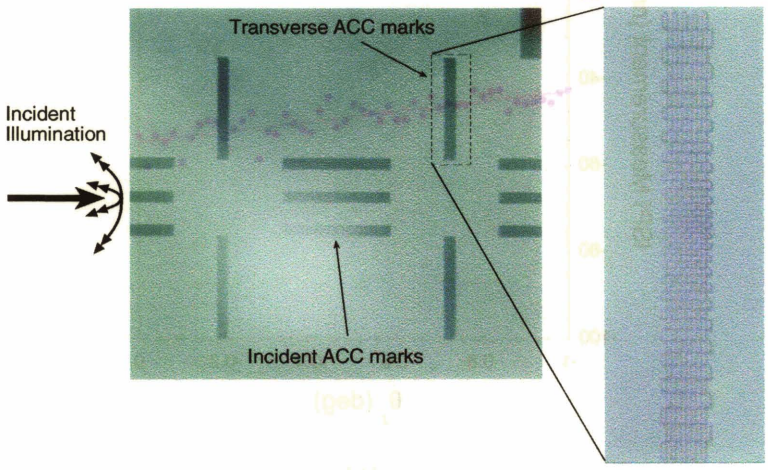
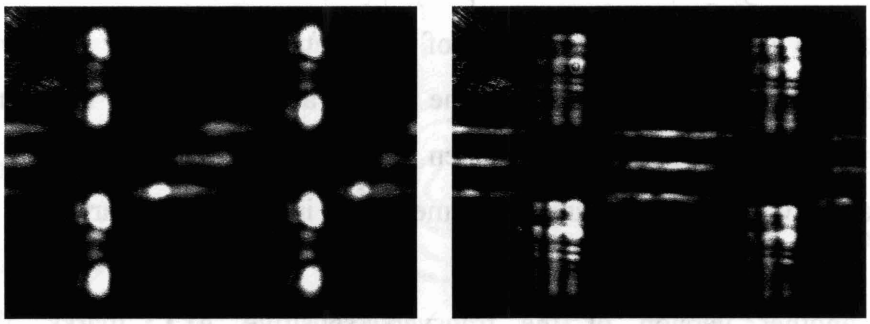
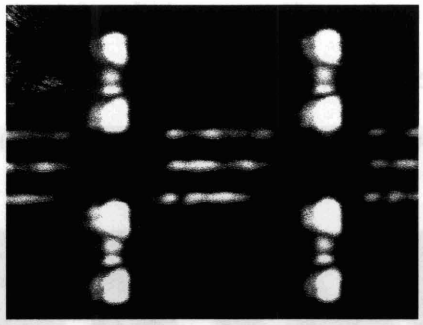


Fig. 7.8. Optical micrographs of angle-checking checkerboard (ACC) marks on a mask. The transverse ACC marks are symmetrically chirped from 2.0 μm to 7.5 μm along the transverse direction, and have a constant 2 μm period in the incident plane. A set of three similar marks, with an inverted center-symmetric chirp in one mark, indicate the incident angle, i.e., the two outer marks are chirped from 2 μm in the center to 7.5 μm at the ends, and the middle mark is chirped from 7.5 μm in the center to 2 μm at the ends.

[27] A phase bias due to angular deviation in the incident plane is not indicated by theory, suggesting a misalignment between the microscope and hexapod coordinate systems, resulting in a slight coupling of transverse-plane rotation bias into the incident-plane rotation.



(a) (b)



(c)

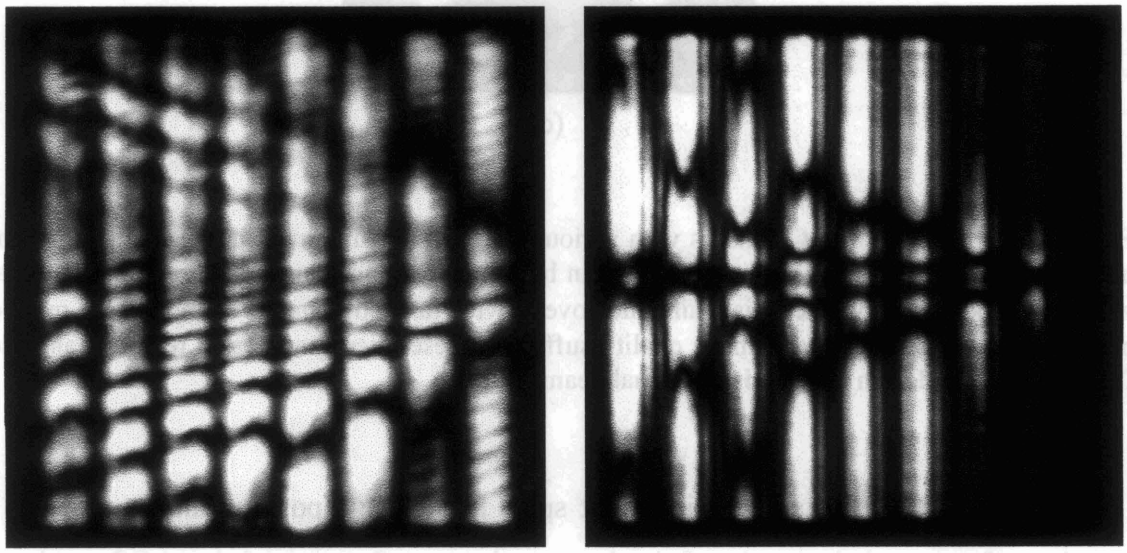
Fig. 7.9. ISPI images of ACC marks with various states of misalignment in the incident (a) and transverse (b) planes. (c) Angular alignment in both incident and transverse directions. The gap was 10 μm, but mark symmetry is maintained over a wide range of gaps, even though the number of diffracted spots will vary. (Image quality suffers in these micrographs as a result of an older ISPI microscope design that used an internal beamsplitter.)

These ACC marks result in bright spots that change position and number as a function of gap and the angle of the incident beam. One set of the ACC marks is sensitive to angles in the incident plane and the other set of ACC marks is sensitive to angles in the transverse plane. Diffraction patterns are formed by the ACC marks in a manner similar to the TCG fringes in Chapter 5. Images from the ACC marks, as formed in the ISPI microscopes, are shown in Fig. 7.9.

If one considers the fringe formation mechanisms, one will see that for the transverse ACC marks, symmetry in the diffracted beams on either side of the incident

plane is imitated by symmetry of the spot patterns. Angles in the incident plane are determined with a related mark consisting of three patterns: two outer symmetric chirped checkerboards with maximum period in the middle and one chirped checkerboard with minimum period in the mark center. When at the proper angle, one spot is seen in the inner mark and two spots are arranged symmetrically in each outer mark.

In another version of the transverse-sensitive ACC marks, the chirped checkerboards have a 1.0 μm period in the X direction, and linear chirps in Y with a center period of 1.00 μm, and maximum periods of 1.25, 1.50, 2.00, 3.00, 4.00, 5.00, 10.00, and 20.00 μm, respectively. The dimension of each checkerboard is 20x400 μm. Images of these marks, as seen in an ISPI microscope, are in Fig. 7.10.



(a) (b)

Fig. 7.10. CCD images of the ACC marks at transverse angles, viewed with an ISPI microscope. These ACC marks contain a range of center-symmetric chirped checkerboards with periods varying from a center period of 1.00 mm to end-periods varying between 1.25 mm and 20 mm, with a 1.0 mm period in the Littrow direction. Mark lengths are 400 mm. (a) Angular misalignment is indicated by significant asymmetry in the fringe patterns, compared with a horizontal centerline. (b) ACC fringes corresponding to maximum alignment, indicating the diffracted beams have symmetry on both sides of the incident plane. Sensitivity of these ACC marks to angular variations in the transverse plane is <0.01°, indicating suppression of error in ISPI fringes to <0.12 nm. (Micrographs taken using a newer ISPI microscope without an internal beamsplitter.)

Observing the ACC marks in Fig. 7.10 while adjusting the hexapod stage indicates that the sensitivity of the ACC marks is beyond the resolution of the hexapod: the angular resolution of the ACC marks is < 0.01°, implying that the ISPI phase bias can be reduced to < 0.12 nm.

7.4 Axial Rotation of the Camera

The standard ISPI phase algorithm detects the phase difference between the average of position of two sets of fringes. Each of the two fringe sets is averaged along pixels in the X direction prior to analysis with a cross-power spectrum. Averaging presupposes that the fringe position is aligned to the X-coordinate of the image. Misalignment of the coordinate systems of the ISPI marks and the CCD camera will result in a perceptual error in the spatial phase relation, i.e., an unintended phase bias due to opposite shifts in the mean position of inclined fringes on the two sides of the mark. The amount of the phase bias Φ_B is expressed as

$$\Phi_B = \tan^{-1}\left(\frac{W}{2P_F}\right) \tag{7.1}$$

where W is the width of the mark (including both fringe sets), and P_F is the fringe period. Three methods of correcting this error have been implemented.

The most straightforward way to avoid errors introduced by camera rotation is to align the X-axis of the camera to long, narrow diffraction gratings on the mask or substrate. Two parallel diffraction gratings are used, with 1 μm period in the incident plane, 5 μm width, and a separation of 30 μm in the transverse direction. When viewed with an ISPI microscope, an electronic fiducial line is superposed on the image, and the camera is mechanically rotated to align the diffracting bars on the mask and substrate to

the electronic fiducial line. The diffracting bars extend throughout the horizontal field of view of the camera (800 μm), and easily could be included in the streets between dies. Let us assume that the parallelism of the electronic fiducial line and the bars can be detected within 1 pixel (7.4 μm), p_1 = 2.00 μm, p_2 = 2.05 μm, and the fringe period P_F = 41 μm, with a mark width of 400 μm and separation of 200 μm between the centers of each fringe set. By similar triangles the best rotational alignment results in a fringe displacement of Δy_F = 200(7.4/800) μm = 1.85 μm. The relative phase error is 1.85/41 μm = 0.0451x2π. One fringe cycle corresponds to a displacement of ~0.5 μm, hence the unintended phase bias is 0.5μmx0.0451 = 23 nm. Clearly, an additional means of removing phase bias is desirable.

There are two alternatives: in one case we can easily modify the algorithm, by splitting the analysis of each fringe set into two or more parts, and calculating the rotation angle from these parts. (In the limit, we can calculate the phase of each column in the image, compared to an initial column. We have used this algorithm at times, but it is not commonly used due to the large number of FFTs.) In the other case, we can add an additional set of fringes, as shown in Fig. 7.11.

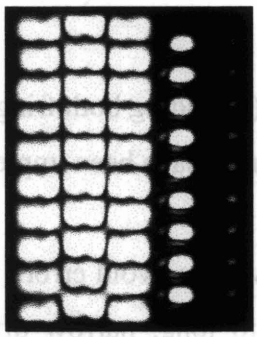


Fig. 7.11. ISPI image of a 3-part mark. Fringe periods result from {p_1, p_2}, {p_2, p_1}, {p_1, p_2} grating arrangements, where the superposed pairing indicates {mask period, substrate period}. Mask and substrate bar arrays are seen to the right of the fringes. Errors due to rotation of the camera are eliminated automatically by the difference in phase measurements between (i) the phase between the first and third fringe sets, and (ii) the phase between the second fringe set and either of the outer fringe sets. This algorithm was suggested by Dr. Patrick N. Everett. The same technique can be applied to gapping fringes.

The third fringe set has the same grating periods as the first set, and there will be a zero phase difference between first and third fringe sets if the camera rotation is zero. Finite camera rotation is eliminated by subtracting the phase between the first and second (counterpropagating) fringe sets from the phase between the first and third (copropagating) fringe sets. This scheme is limited by the rotation of the fringes outside of rectangular regions of interest in the image and rotation-induced phase ambiguity.

The advantage of this scheme is that only two phase measurements are required. Another advantage is that the marks are symmetric with respect to the centerline, which may make the marks insensitive to flips or rotations from printing on transparent surfaces. The disadvantage is that an additional set of fringes is added, which will either require more area, or narrower fringe sets in the same area.

7.5 Microscope Variation

High measurement sensitivity is meaningless unless it is reproducible with multiple instruments. This is especially significant in ISPI, which uses three microscopes to determine three-dimensional position. To demonstrate reproducibility, we arranged two ISPI microscopes to observe the same alignment mark from opposite directions. The purpose of this experiment is twofold: first, it will reveal any consistent measurement bias and determine if small deviations in detected alignment are real, or noise. This experiment also will show any variation of the alignment signal from one microscope to another that may be caused by differences in the optics, CCD cameras, fibers, mirrors, microscope alignment, etc.

In the experiment, two ISPI microscopes, looking from opposite directions, observed the same alignment mark. Each microscope mirror was adjusted to the optimal angle based upon observations of the ACC marks (Section 7.3). Axial camera rotations

were minimized by physically rotating the cameras to match fiducial lines on the mask, and using multiple analysis regions on each fringe set. Illumination from one microscope produces a specular reflection off the mask that is directed into the opposite microscope. As a result, the measurements were taken alternately, with the light from one microscope blocked while measurements were made through the other microscope. Sequences of such alternating measurements were made, for a total of 150 measurements through each microscope. The results are shown in Fig. 7.12.

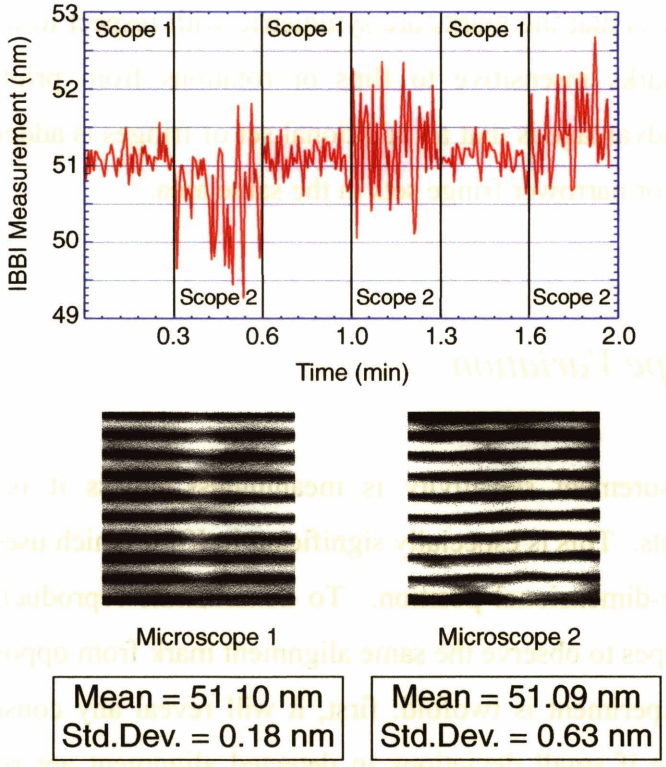


Fig. 7.12: Two ISPI microscopes alternately observe the same alignment mark from opposite directions, with the mask locked to the substrate at a 3 μm gap. The difference in mean values measured between the two microscopes is 0.01 nm.

The mask was mechanically locked to the substrate by 3 μm-tall aluminum spacers (a.k.a. studs) evaporated on the mesa rim surrounding the mask membrane. Setting the mask on the studs implies that the mask and substrate are held in mechanical

contact by friction, thereby avoiding relative motion. Any effects of vibration or drift of the mask stage are eliminated when the mask is in contact with the substrate through the studs.

Both microscopes measured an alignment offset of ~51 nm, with a standard deviation <1 nm, and a difference in mean values of 0.01 nm. Microscope #2 gave noisier fringes than #1. However, the noisy fringes did not adversely affect the measurement significantly, indicating the robustness of the measurement.

7.6 Summary

One of the distinctive advantages of ISPI over other interferometers is the ability to analyze two displacement-sensitive signals at once, with the same sensor, and obtain a position measurement from the relative phase of the two measurements. The advantage of this method is that whatever affects one measurement will also affect the other, adjacent measurement. The two fringe patterns are observed side-by-side, in close proximity, and whatever affects one measurement will usually affect the other measurement in the same way, and by the same amount. We refer to this property as "mutual phase referencing".

In Section 3.2.2, we calculated that resist coverings will change the optical path lengths, but will not change the angles of diffraction outside the resist or the phase of the interference fringes. In this chapter we described experimental evidence of the immunity of ISPI to layers covering the substrate marks. The intensity of the beams passing through resist was diminished, having a slight effect on the image contrast, but essentially zero variation was observed in the spatial phase. Similar results are found for layers of polysilicon and aluminum.

Section 7.2 examined the effects of grating defects on the ISPI alignment signal. Shifts of a single grating line were essentially unobservable, as were omissions of grating lines amounting to ~10% of the fringe period. The fringes also allowed wide latitude in grating line-space ratio.

Section 7.3 described the effect of viewing-angle rotations on the fringe phase. Experiments indicated a phase bias of ~12 nm/deg. New interferometric marks were introduced to detect viewing angles within an ISPI image to <0.01°, sufficient for <0.12 nm phase bias.

Section 7.4 considered the problem of phase bias introduced by rotation of the camera around the optical axis. Coarse and fine methods of camera-fringe alignment were discussed, including a three-fringe variation on the ISPI mark that automatically cancels rotation.

Section 7.5 responds to the necessity of measurement reproducibility in different microscopes. Two ISPI microscopes measure fringes from a single alignment mark, and report mean alignment consistency to 0.01 nm.

8 ISPI Alignment System

This chapter will depict the ISPI aligning and gapping apparatus. The imaging system consists of three ISPI microscopes, four diode lasers and a spectrometer. Positioning stages utilize a custom 5-axis piezo stage with 0.5 nm X and Y step size. An x-ray source is used for aligned exposures. A procedure for x-ray source alignment completes the chapter.

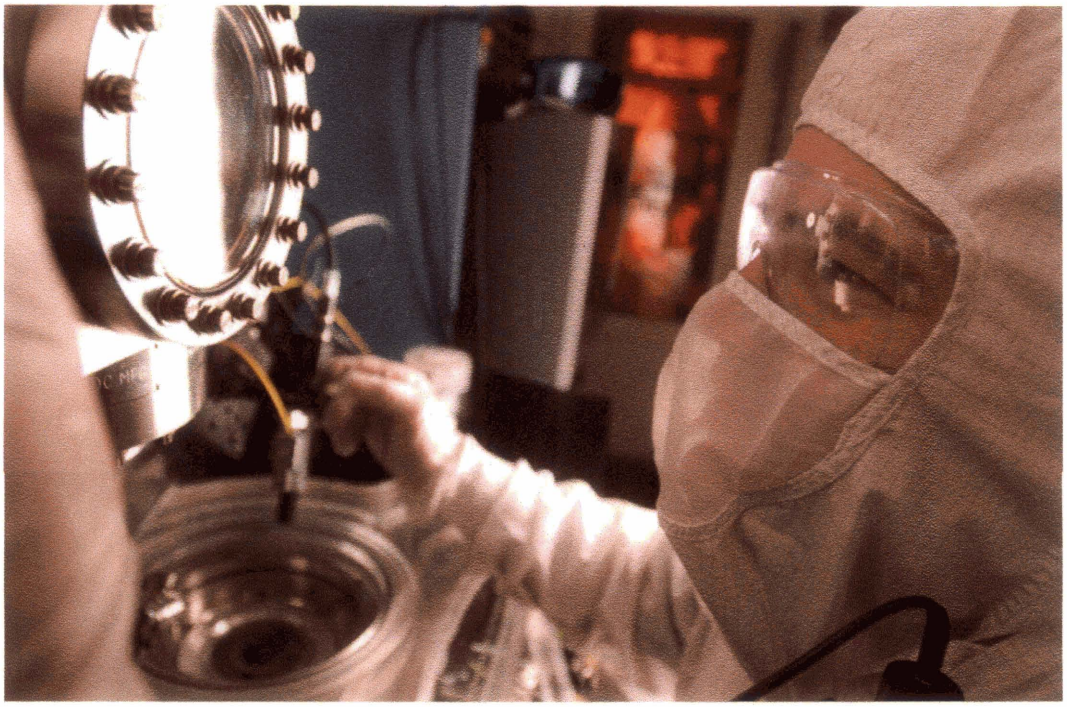


Fig. 8.1. ISPI alignment system and user
(MIT Research Laboratory of Electronics Calendar, June 2004).

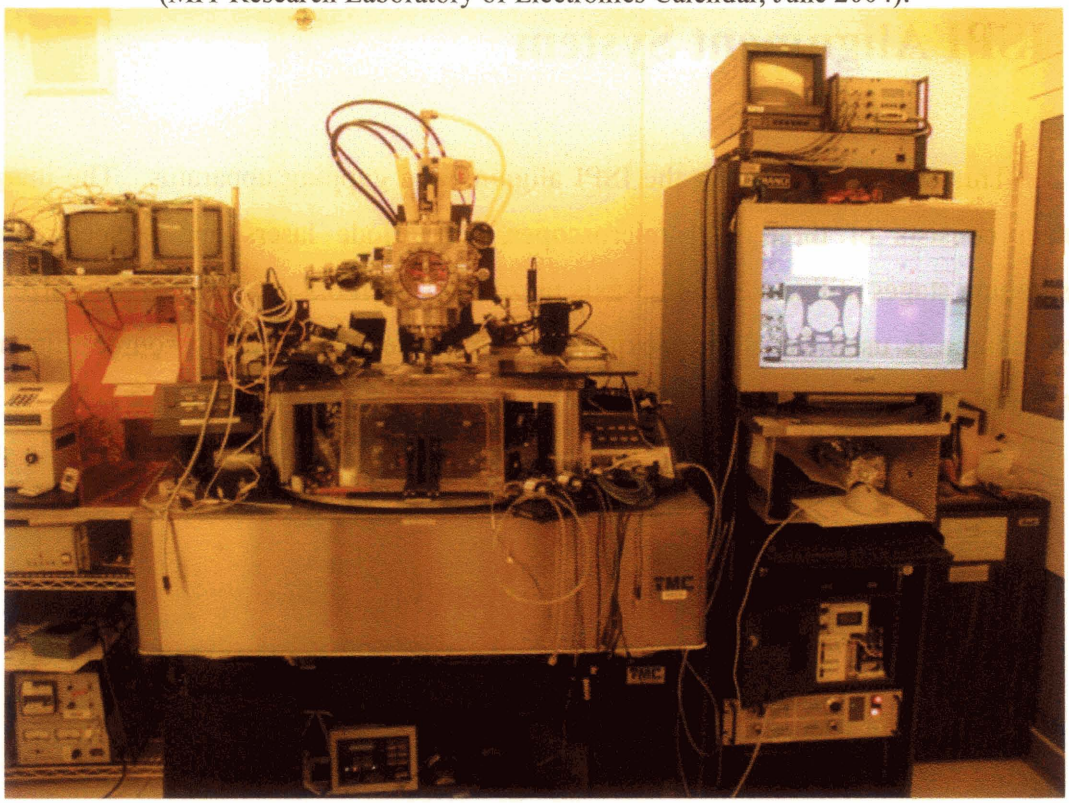


Fig. 8.2. ISPI nanopositioning apparatus and control system.

8.1 Mask Stage

The mask stage consists of a PI P-730.20 XY piezo stage with integrated capacitive sensors. The x-ray mask is held in an annular clamp. The clamp rests in an outer annular ring, attached to the XY stage. The mask is held by gravity in a bevel in the lower ring, but is free to move in Z when the mask is brought into contact, thereby releasing the mask from the outer ring. Due to the height of the clamp ring, the stages can continue to move towards the substrate after mask-substrate contact, in which case the mask is resting on aluminum studs deposited on the circumference of the mask mesa, and is independent of any stage drift. The mask is held in place on the substrate by its own weight. An E-710.4CL four-channel digital controller with 32-bit resolution in the piezo amplifier controls the XY piezo mask stage. The range of the XY piezo stage is 50 μm and the resolution is 0.5 nm.

The XY piezo stage is held in a custom Z-tip-tilt stage with three PI P-841.1B piezos with integrated strain gauges. The piezos are actuated by an E-503 LVPZT amplifier, and the displacement is linearized over the full range of travel using an E509-S3 PZT servo controller. An E-516 GPIB interface module communicates with the alignment system's main computer. The range in each Z piezo is 15 μm and the resolution is 1 nm.

The XY piezo stage and the three Z piezos are controlled by the main system computer through a National Instruments GPIB interface card.

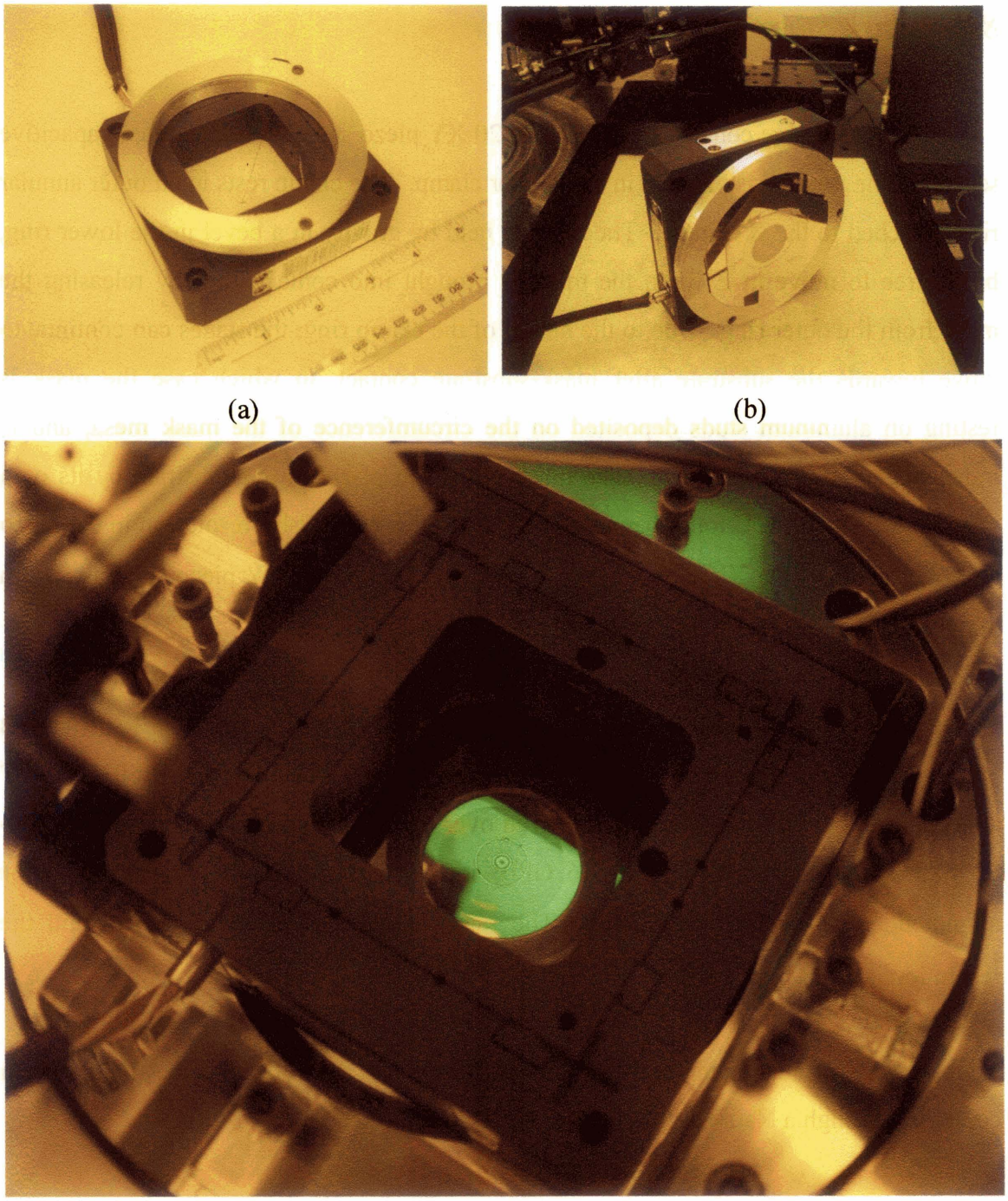


Fig. 8.3. (a) Photograph of the XY piezo stage for the mask. (b) An annular ring holds the mask. The ring is contained between circular plates attached to the XY piezo stage. (c) Photograph of the custom Z-tip-tilt piezo stage in the alignment system, with the XY mask stage and an x-ray mask in the exposure position. The XY stage is inverted when used in the alignment system. The XY piezo stage is held in the system by a three-piezo Z-tip-tilt stage, which in turn is held in a DC-servo-driven 25 mm-travel-range Z-tip-tilt stage.

8.2 Substrate Stage

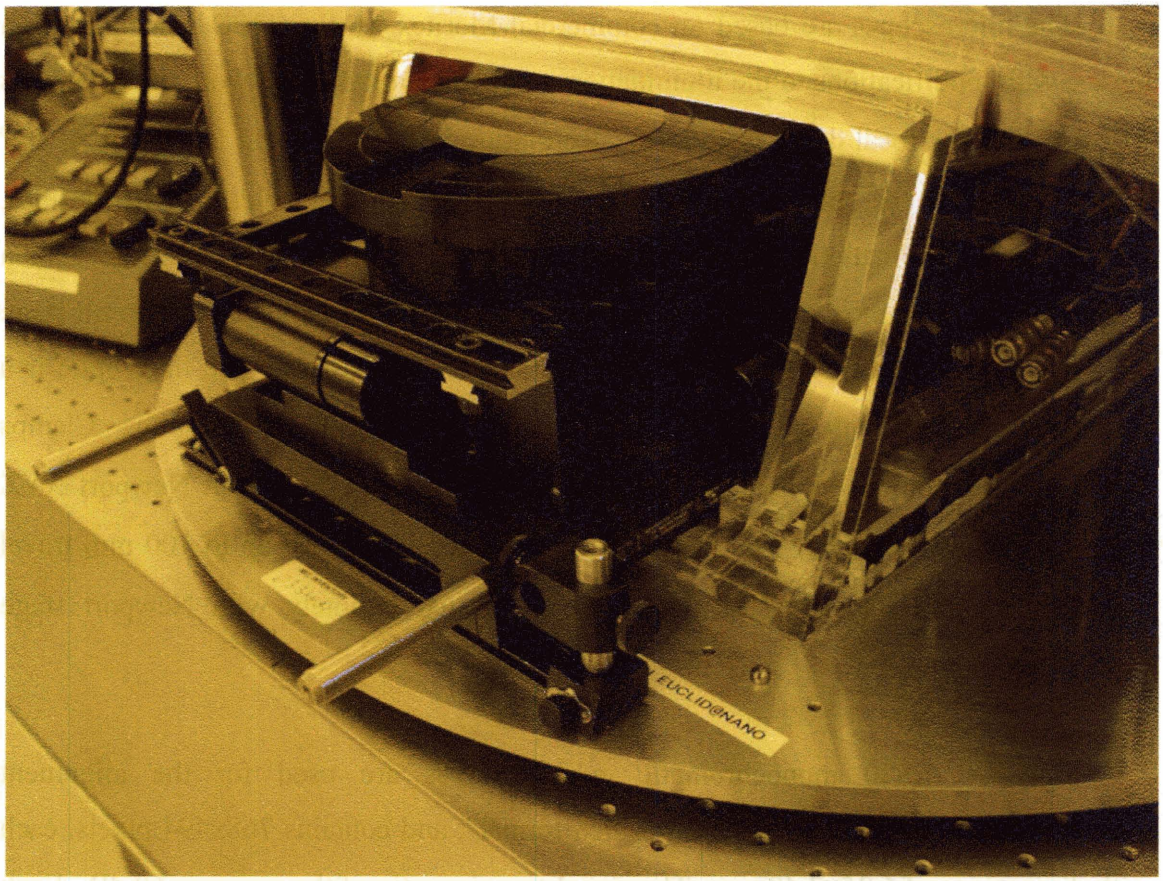


Fig. 8.4. Photograph of the substrate stage in the sample-loading position. The stage contains X, Y, and rotation DC drives with integrated optical encoders. X and Y range is 100 mm, and minimum step size is 25 nm. Rotation range is 360 degrees and step size is 0.5 arcsec.

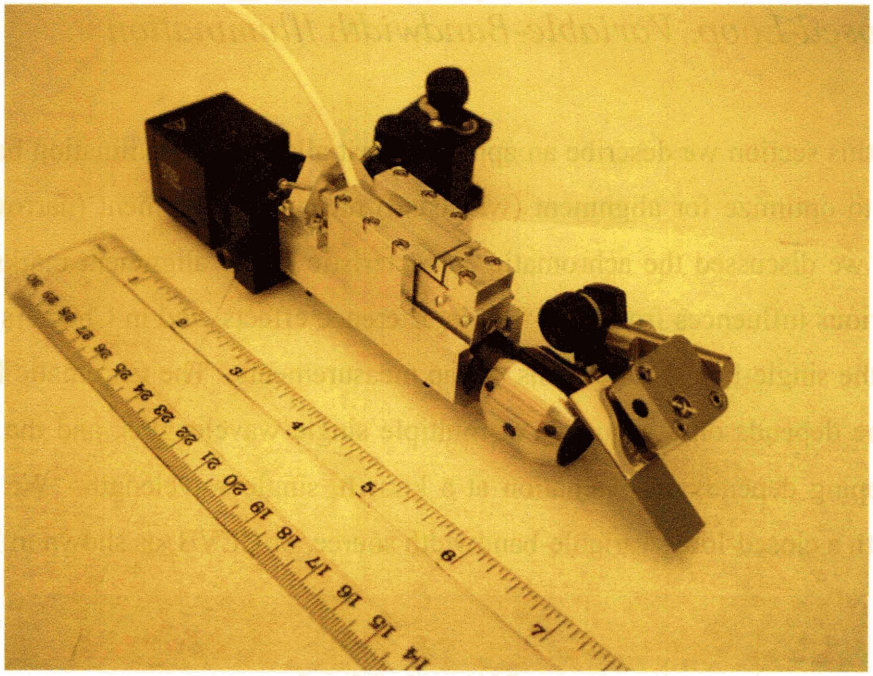
The substrate is actuated with a Newport PM-500 XYθ stage. The X and Y axes employ DC drives with optical encoders, with a minimum step size of 25 nm in X and Y. The rotation stage has a minimum step size of 0.5 arcseconds. The three axes of the substrate stage, as well as three DC drives used for coarse gap control, are controlled with a Newport PM500-C 6-axis stage controller. A GPIB card interfaces to the main system computer. The substrate stage holds samples up to 200 mm diameter.

8.3 *Imaging System*

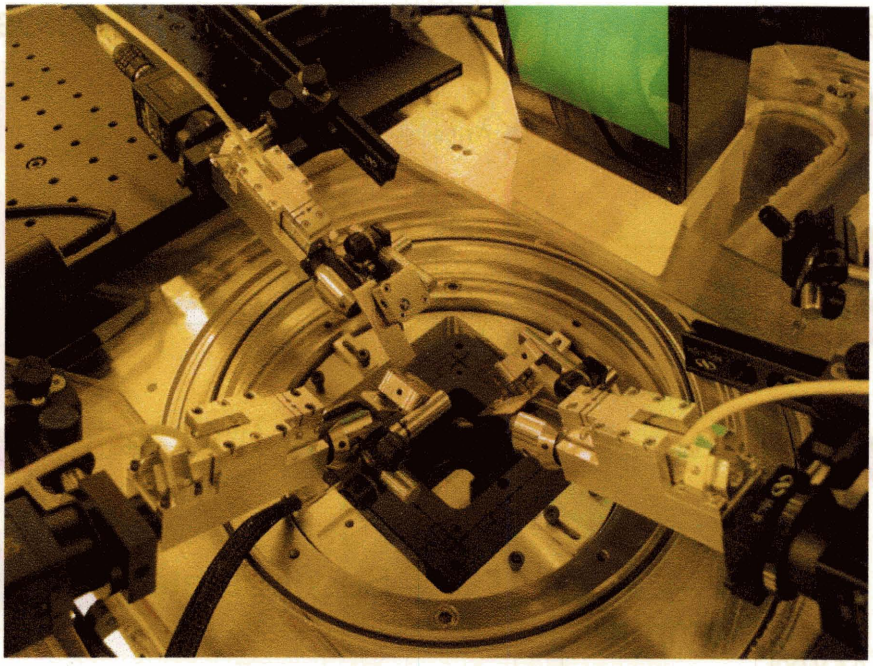
The microscopes are based on commercial Moritex MML6-110D machine vision lenses, with modifications to the illumination input. The microscope working distance is 110 mm, and the focal depth is 150 μm. Resolution is 11 μm. Light is incident at a 4° angle from above the optical axis using a custom designed fiber holder [PNEC design]. Imaging and illumination paths are deflected towards the mask and substrate with a silicon mirror. The mirror angles are adjusted with a New Focus kinematic tip-tilt stage and Newport microrail components, and a custom rail attachment. The fiber optic is a single-mode design, with a 3 μm core and a pigtail collimator, from OZ Optics. Clamps and microrail components hold the microscope body to an independent Newport XYZ stage for each microscope. DC motors with 1 μm resolution and up to 100 mm travel range actuate the stages. The motors are controlled remotely with Newport stage controllers.

Sony XC-ES30 monochromatic cube cameras are used for the alignment experiments. The image sensor is "1/3-inch format", and contains 768x494 pixels, each with dimensions of 6.35x7.40 μm. The signal-to-noise ratio of the sensor is 60 dB. Each camera uses a single Hirose cable to supply power and contain the video output.

The analog video streams from the microscopes are digitized using National Instruments PCI-1408 framegrabbers. Each framegrabber can acquire images from up to four inputs. Two framegrabbers are used in a rotating framegrabbing order, to avoid video glitches when changing input channels. The framegrabbers can be synchronized for simultaneous image grabs. LabVIEW IMAQ software is used for image acquisition and analysis.



(a)



(b)

Fig. 8.5. (a) Photograph of an ISPI dual-path microscope. (b) Photograph of three ISPI microscopes obliquely viewing ISPI marks on a mask. Illumination is brought in with single-mode fiber optics, collimated, and reflected by a metalized silicon substrate sliver down to the mask and substrate at a near-Littrow angle. Diffracted beams return at a similar angle, and are imaged onto CCD cameras by 6x microscopes at a working distance of 110 mm from the mask.

8.4 *Closed-Loop, Variable-Bandwidth Illumination*

In this section we describe an apparatus that allows the illumination bandwidth to be varied to optimize for alignment (wideband) or gap measurement (narrowband). In Chapter 3 we discussed the achromatic characteristic of the alignment marks that helps avoid spurious influences from thin-film interference effects, and in Chapters 4, 5, and 6 we noted the single-line requirements in gap measurements. The Chromatic Fabry-Perot gap scheme depends on a sequence of multiple single wavelengths, and the Transverse Chirp Gapping depends on evaluation at a known, single wavelength. We meet these criteria with a closed-loop, variable-bandwidth source, or CLVB, as shown in Fig. 8.6.

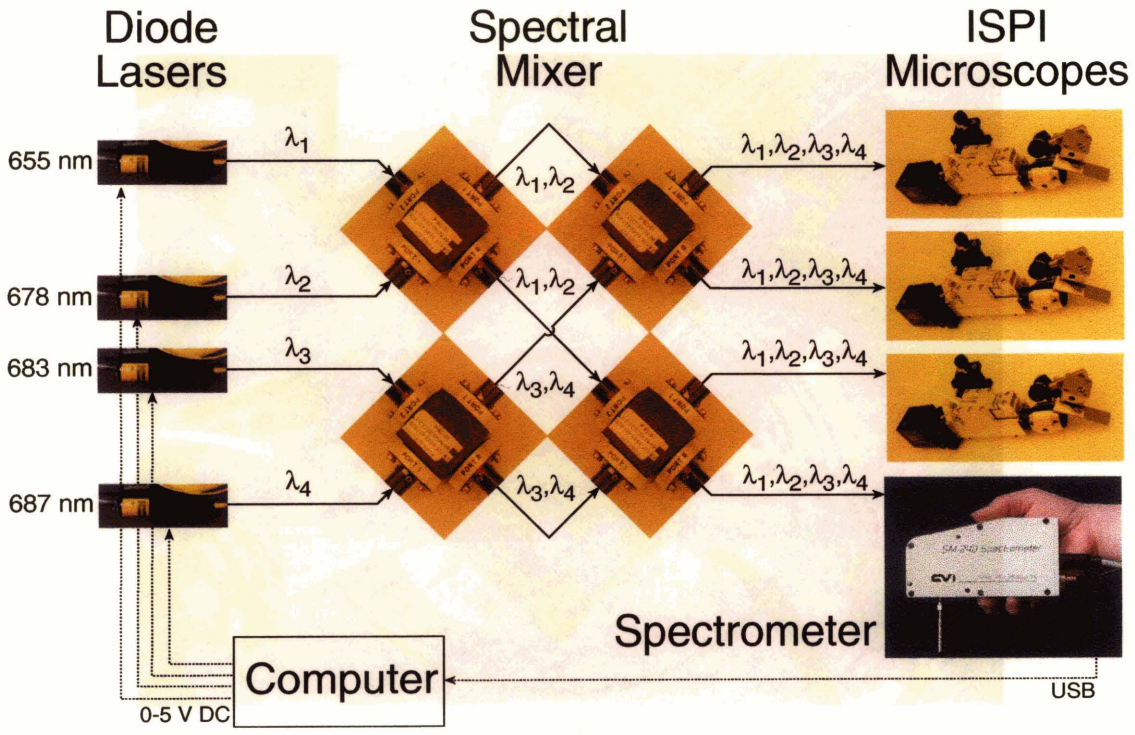


Fig. 8.6. Schematic of the Closed-Loop, Variable-Bandwidth (CLVB) light source. The CLVB source optimizes the bandwidth for either aligning (broadband, equal power), or gapping (single line). (Images of components are to different scales.) A set of four fiber-coupled beamsplitters mixes the four input wavelengths equally in four outputs. ISPI microscopes use three outputs, and the fourth is used for spectral analysis.

The CLVB light source consists of four Lasiris diode lasers (655, 678, 683, and 687 nm, 30 mW apiece) whose spectrum is mixed in a combination of four beamsplitters from OZ Optics. The beamsplitters are cross-coupled to achieve equal amounts of each spectral line in each of four fiber outputs. The ISPI microscopes use three of these outputs, and the fourth output is used by a compact spectrometer (CVI Optics), whose output is transferred to a computer and used to ascertain the wavelength and power transmitted through the spectral mixer. The power of each laser diode is subsequently adjusted to obtain the desired spectral content.

8.5 Source Axis Alignment

When a point x-ray source is employed, such as the electron-bombardment x-ray source used in this dissertation, the position of the source relative to the mask patterns must be known prior to exposure to ensure pattern placement accuracy. At a finite gap, exposed features will display a geometric offset proportional to the gap and the radial distance from a point of symmetry directly underneath the source. The point of symmetry intersects a ray from the source that is normal to the substrate surface. This ray defines an axis between the source and mask, which we refer to as the *source axis.*

Even at the large source-mask distance (274.5 mm) and small mask–substrate gaps we typically use (1 to 5 μm), image magnification is not negligible, and in fact may amount to several nanometers. Specifically, for nanometer overlay the gap must be known and controlled to <0.1 μm and the source axis must be known within about 100 μm.

Locating the source axis may seem an onerous restriction for large-area patterning, but setting the gap to a known value, and biasing the feature positions in the mask pattern can compensate for the source-divergent magnification. The trick is to find

the source axis with sufficient resolution to reduce the source-divergence errors to <1 nm[28].

In the x-ray source attached to the ISPI apparatus (electron-bombardment source with a copper target, with a peak output at λ = 1.3 nm), a 1-μm-thick, 1-cm-diameter silicon-nitride membrane forms a lower part of the vacuum chamber containing the x-ray source. Although hanging a plum line would be a very simple means of finding the source axis, it is not possible to do so through the membrane vacuum window. Measuring the position of the external vacuum chamber is also unsatisfactory, due to thermal deformation of the electron filament that causes the focal spot of electrons on the copper target to migrate (with a slow time constant) around the geometric center of the target. Laser positioning is another possibility: a laser beam could be split, with one arm retroreflected from a fiducial point on the substrate and the other arm centered on the copper target. Coincidence of both beams on the target would signify alignment of the source. However, this too is insufficient due to uncertainty in the exact position on the target from which the x-rays emanate. The solution we pursue in this section is to measure the x-ray flux directly, and from its divergence, deduce the source axis. Once the source axis is found, any mark of interest on the mask or substrate is aligned to the source axis with an optical microscope, which is sufficient for source-axis alignment to <100 μm.

[28] A commercial x-ray lithography system would use a collimated beam, such as in a synchrotron or collimated-plasma x-ray source. Other forms of proximity lithography would not have this problem. For example, in imprint lithography there is no need to have a collimated UV source, since the template is in contact with the organic monomer on the substrate.

8.5.1 Scanning X-ray Pinhole Camera

The first step in finding the x-ray source axis is coarse alignment to position the source at the center of numerous apertures along the x-ray beampath. The critical apertures include the electron shield, which is about 1 cm above the x-ray source, the 1-micron thick, 1-cm-diameter silicon-nitride vacuum window, the stainless steel bellows connecting the vacuum chamber to the helium box, and the viewport aperture. The viewport aperture consists of an 18.75 mm diameter hole bored through a 12.5 mm thick, 100 mm diameter optical flat in the top of the helium chamber. This viewport allows multiple ISPI microscopes to view any portion of the x-ray mask during exposure. The viewport is in a fixed position in the center of a 12" diameter stainless steel plate that forms the top of the helium chamber. Since the viewport is the aperture furthest from the x-ray source, it has the greatest potential to obscure x-rays at the mask plane.

Adjustments in position are made to the vacuum chamber containing the x-ray source. DC motors on 25 mm cross-roller bearings control the x-ray source position in X and Z. Y adjustment is made manually, by releasing a clamp containing the neck of the ion pump, and sliding the neck through its yoke. Rotational adjustments in the X-Z plane are made manually in the same fashion.

To perform the coarse alignment we use a commercial board-level CCD camera [8.11] with an opaque disk in the camera's C-mount containing a 1-mm-diameter pinhole, mounted to the rear of the substrate stage. The CCD has its cover glass removed to allow penetration of x-rays. The substrate stage is displaced on a cross-roller bearing stage to bring the camera underneath the x-ray source. Custom software scans the substrate stage in a raster pattern in X and Y to generate an image of the x-ray flux passing through the various apertures in the x-ray beampath, building up an image of the x-ray flux at the substrate plane, as shown in Fig. 8.7. At each scan step, the image intensity over the CCD is integrated and displayed as one pixel in a 2D array. The source position is

aligned to the column by iterations of mechanical adjustments and imaging the x-ray flux. The step size selected in a scan was a tradeoff between scan duration and resolution. Initial scans were coarse, with an 8 mm step size, and later scans were at reduced step sizes, down to 500 μm. Finer estimation of the source axis requires another method, described in the following section.

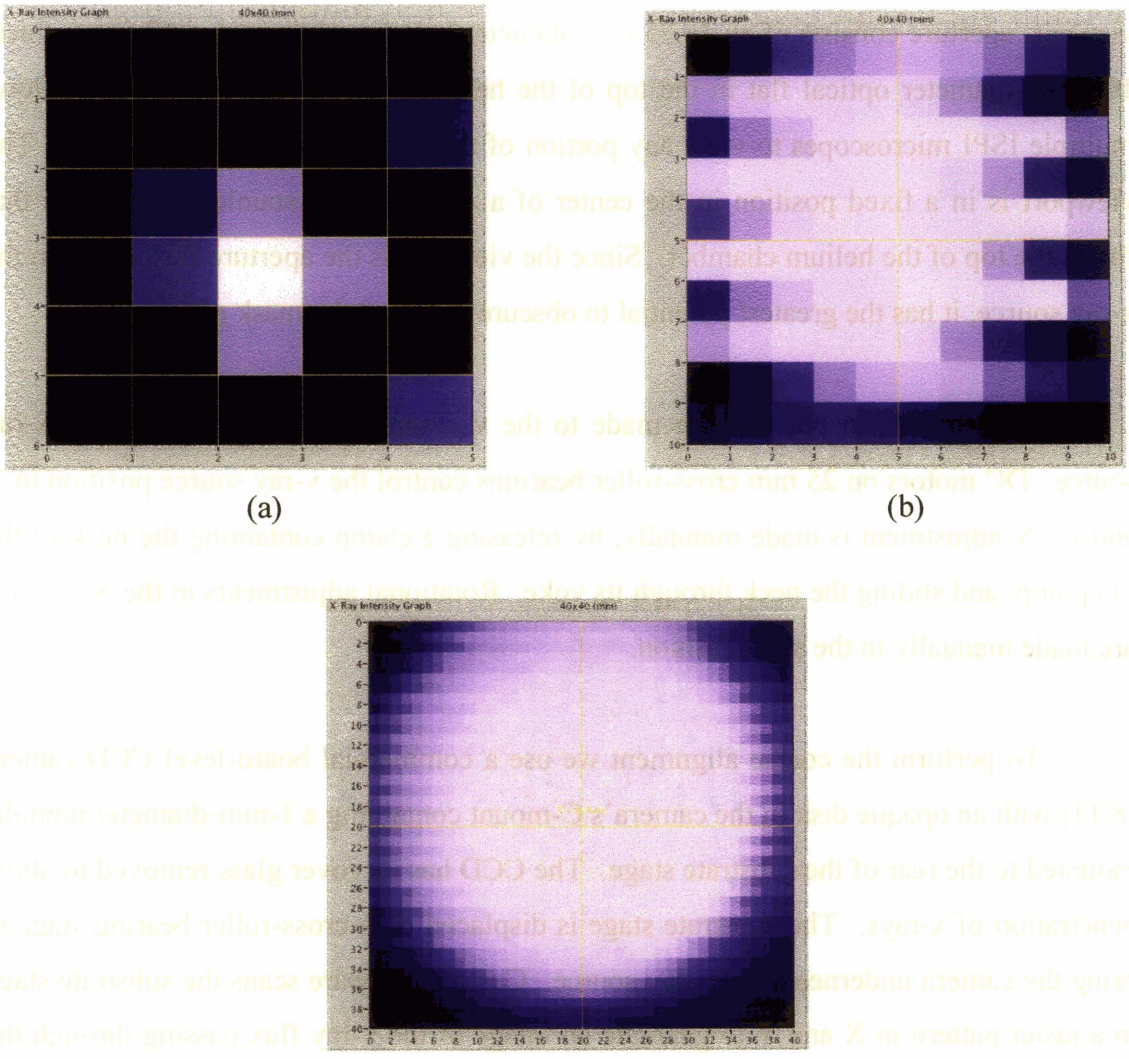


Fig. 8.7. 2D intensity plots of data from the scanning x-ray pinhole camera. The scan is over a range of 40x40 mm, with step sizes of (a) 8, (b) 4 and (c) 1 mm.

8.5.2 *Micropipette Alignment Apparatus*

The apparatus we use to measure the x-ray flux source axis is shown in Fig. 8.8. We use the same CCD camera as in the scanning x-ray pinhole camera. We replace the pinhole with a thin sheet of aluminum-coated mylar (12.7 μm thick) that significantly attenuates optical wavelengths (90% attenuation), but has a negligible attenuation of x-rays. The CCD, with its mylar sheet, is mounted on a tip-tilt stage on the side of the substrate stage. Above the CCD is a close-packed bundle of micropipettes (700 μm inner diameter, 33 mm long, with approximately 30 tubules), mounted on a separate tip-tilt stage.

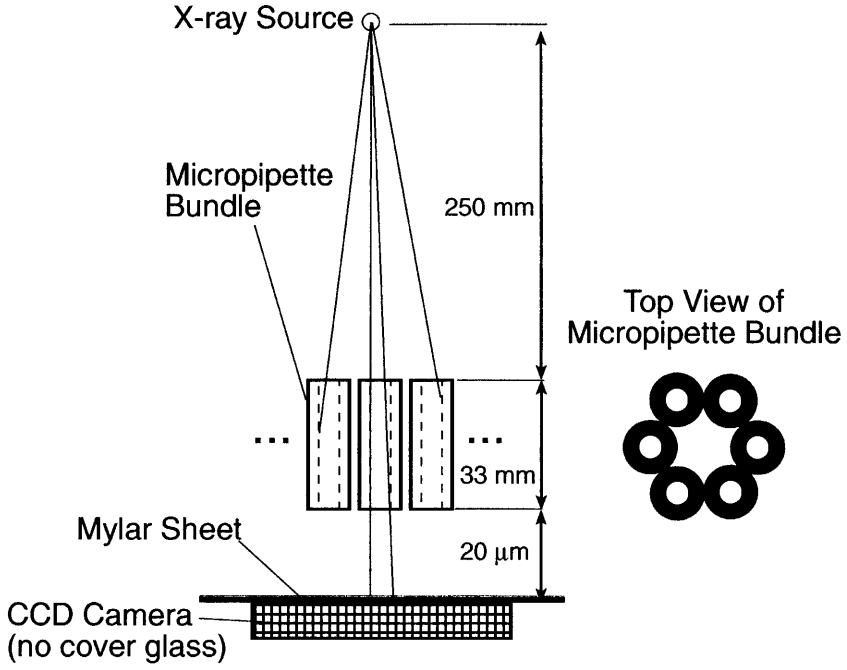


Fig. 8.8. Schematic of x-ray source alignment geometry. Not to scale.

8.5.3 X-ray Alignment Procedure

Prior to finding the source axis, the plane of the CCD is made parallel to the substrate plane using laser retroreflection from both surfaces, and the axis of the micropipettes is adjusted to be parallel to the substrate normal.

The alignment procedure is as follows: a 3x normal-incidence microscope above the substrate contains a beamsplitter behind the objective lens. A diode laser beam is coupled into this beamsplitter, and emerges collimated from the objective, coincident with the optical axis. The laser beam is directed toward the substrate surface with an adjustable mirror. The angles of the mirror are adjusted until the laser beam retroreflected from the substrate surface is superimposed with a second beam resulting from internal backreflection within the beamsplitter, and detected on the microscope's CCD, indicating that the optical axis, as indicated by the laser beam, is perpendicular to the substrate. Since the bright retroreflected beam obscures the weak reflection from the internal beamsplitter, an opaque card intermittently blocks the beam external to the microscope. The relative position of the beams is compared using electronic crosshairs on captured images. (A pinhole in front of the objective lens confirms that the laser is reflected perpendicularly to the substrate.) Next, the substrate stage slides away from the microscope until the x-ray CCD mounted on the rear of the substrate stage is underneath the 3x microscope. The tip-tilt stage holding the x-ray CCD is adjusted to retroreflect the laser beam to the microscope's CCD, thereby ensuring that the x-ray CCD plane is parallel to the substrate. Without moving the microscope, a micropipette bundle is mounted above the x-ray CCD and adjusted on its own tip-tilt stage so that the retroreflected laser beam travels straight down one pipette and returns along the same path. Images of the micropipettes in the alignment procedure are shown in Fig. 8.9. The pipette angles are adjusted until a single optical mode is observed within the pipette (Fig. 8.9(d)), indicating alignment of the micropipettes to the substrate normal within a few microradians.

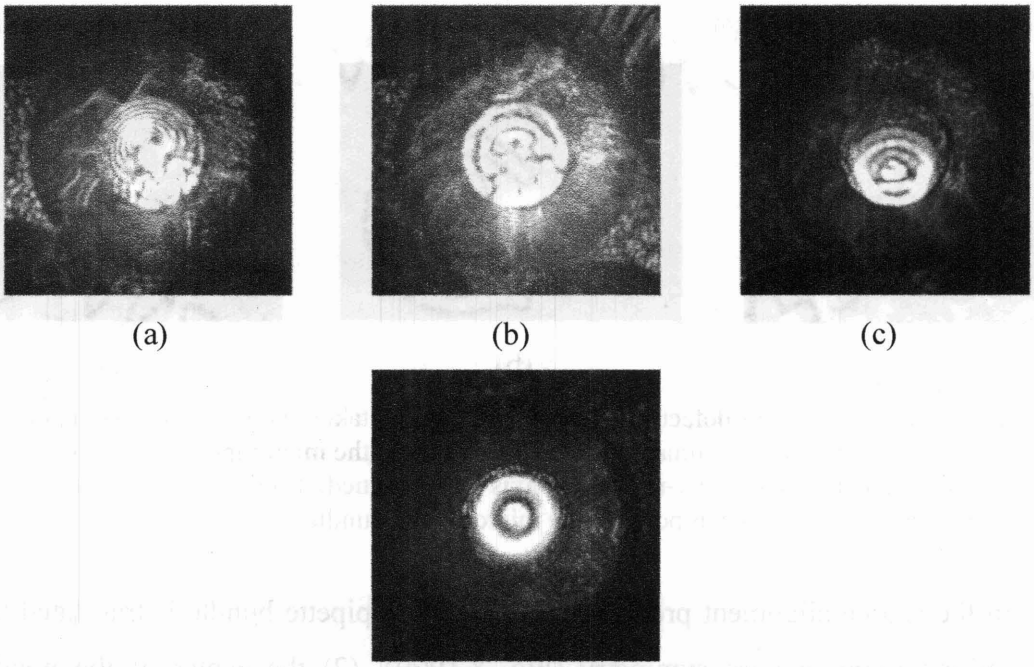
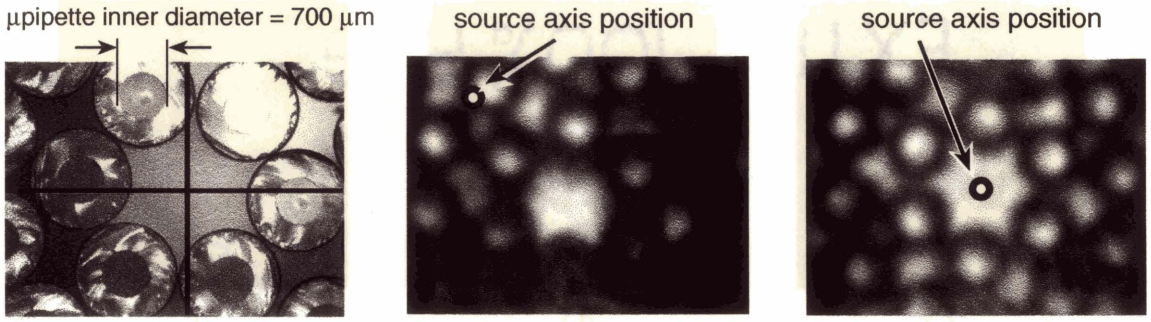


Fig. 8.9. Images of a glass micropipette (inner diameter = 700 μm) taken with a normal-incidence 3x microscope and coaxial laser illumination (λ = 690 nm). The laser is oriented at the normal to the substrate and x-ray CCD. A silicon piece placed over the x-ray CCD camera reflects the laser back through the micropipette. Images (a), (b), and (c) display interference fringes within the micropipette at various tilt angles. Image (d) shows a concentric interference pattern resulting from alignment of the axis of the micropipette to the laser beam normally-reflecting from the CCD.

To determine the source axis, the CCD-micropipette assembly is scanned underneath the x-ray source until (a) x-rays are observed passing through some micropipettes, and (b) the pattern formed by the x-ray flux is symmetric to a fiducial point in the micropipette bundle (i.e., a point defect in the hexagonal close-packed array, Fig. 8.10).

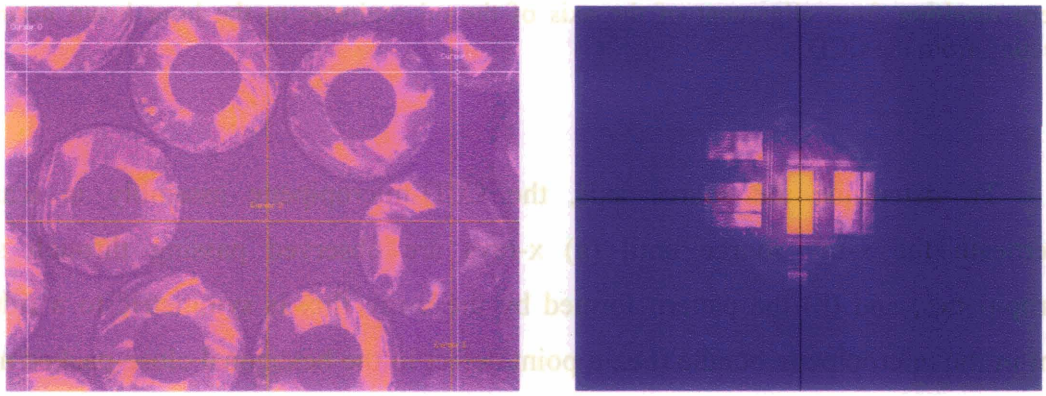
The divergence of the x-ray point source results in obscuration of the x-ray flux by the micropipettes, in proportion to the distance away from the system axis. This radially-dependent obscuration yields a pattern detected by the CCD, shown in Fig. 8.10(b), in which the source axis intercepts the image in its upper left corner.



(a) (b) (c)

Fig. 8.10. (a) Image of point-defect in micropipette bundle, taken with 3x microscope, observing the top plane of the bundle. (b) Image of x-ray flux through the micropipette bundle, observed by the x-ray CCD camera. Here the x-ray source axis is misaligned. (c) Image of x-ray flux with the source axis aligned to the fiducial point of the micropipette bundle.

In the source alignment procedure: (1) the micropipette bundle is translated to the point of maximum pattern symmetry (Fig. 8.10(c)), (2) the center of the bundle is observed with the 3x optical microscope and denoted with crosshairs (Fig. 8.11(a)), and (3) a point on the x-ray mask, observed with the microscope, is aligned to the crosshairs (Fig. 8.11(b)).



(a) (b)

Fig. 8.11. Images acquired with a normal-incidence 3x microscope. (a) After the micropipette bundle is centered on the x-ray flux, the 3x microscope views the top of the bundle, and electronic crosshairs denote the center of the bundle. (b) The mask is loaded, and the desired mark, as viewed on the mask, is aligned to the crosshairs.

Using a 3 μm gap and a source-mask separation of 274.6 mm this scheme enables alignment of the mask to the system axis within 100 μm, corresponding to overlay runout <1 nm.

8.6 Summary

This chapter described the components of the nanopositioning apparatus, including a five-axis, 0.5 nm-resolution piezo mask stage, ISPI microscopes and closed-loop variable-bandwidth illumination system, and a method of aligning the x-ray source to a fiducial point on the mask that directly uses the x-rays emanating from the point source to locate the source axis. We estimate that we can align the source axis and the alignment mark to <100 microns, corresponding to <1 nm error with a 3 μm gap.

9 Lithographic Overlay Experiments

In this chapter experimental evidence demonstrates sub-3 nm overlay and sub-1 nm detectivity, using ISPI metrology. Experiments are done with x-ray exposures of a mask and substrate, and the veracity of the ISPI measurements is determined by the difference between fringe measurements in ISPI microscopes and phase measurements of patterns in exposed resist. Overlay experiments are performed with the mask in contact with the substrate, as well as with feedback control of the mask during exposures at finite gaps. Alignment of upper and lower gates in a double-gate MOSFET transistor is demonstrated using multiple ISPI feedback-controlled exposures.

9.1 *Alignment Detectivity Experiment with Contact Exposure*

An experimental verification of the detectivity of ISPI compared the spatial phase detected by ISPI microscopes prior to exposure with the spatial phase of a moiré pattern in the resist after exposure and development. Before the exposure, a fringe pattern resulting from the interference of diffracted beams was observed by the ISPI microscopes. After the exposure, a moiré pattern between the patterned resist and the complementary pattern etched into the substrate was viewed at normal incidence by an ordinary optical microscope. Despite the fact that the two types of fringes are observed with different microscopes with different viewing orientations, and are formed by different physical mechanisms, the spatial phase should be the same in both cases. The degree of agreement between the two moiré patterns is taken as an indication that the pattern transferred to the resist is highly correlated with the ISPI measurement and the ability of the fringes seen before exposure to successfully predict the critical figure of merit: the phase, and hence alignment, of patterns exposed in resist. These experiments were done with the mask in contact with the substrate to (a) avoid drift of alignment during exposure and (b) to eliminate alignment artifacts due to the divergence of the x-ray point source.

In preparation for this experiment, numerous factors required consideration. First, an array of marks was written on a mask to allow statistics to be generated using a single mask and exposure. Multiple marks, distributed through different e-beam fields, raised issues of pattern placement accuracy. Interfield errors, such as stitching errors, exist between e-beam fields (409.6 x 409.6 μm per field), and constitute unpredictable shifting, rotation and scaling of the field by 10s of nanometers[29]. Interfield errors, if included in the detectivity of ISPI, would invalidate the experiment. Marks written within a single

[29] The problem of e-beam pattern placement accuracy can be solved by spatial-phase locking methods [9.1]. But using Spatial Phase-Locked E-Beam Lithography (SPLEBL) to write the mask would only convolve its performance with that of ISPI, and would not truly elucidate ISPI detectivity.

field, i.e., intrafield distortions, however, can contain distortions below 1 nm. The dimensional consistency from mark to mark may be further complicated by distortions in the mask resulting from absorber stress, as well as stress induced by the mask holder. Pattern magnification due to the exposure point source was eliminated by performing the exposure with the mask in intimate contact with the substrate.

The question of ISPI detectivity was separated from these issues simply by taking the difference of the spatial phase measurements before exposure (in an ISPI microscope) and after exposure (in resist, viewed with a normal-incidence microscope). Of course, the absolute value of the measurements exhibited the distortions discussed above, but the difference of the spatial phases is the only quantity necessary to determine correlation between observed and actual alignment.

A final issue is how to measure with nanometer accuracy the placement of the patterns in the resist relative to patterns etched previously in the substrate. Since we are interested in nanometer or sub-nanometer measurements, there is no measurement method currently available in the semiconductor industry suitable for this task. However, moiré methods are well known to yield sufficiently high sensitivity, and fortuitously, a moiré pattern is formed between the resist and the complementary grating mark etched into the substrate, which has the same fringe period as the interference fringes viewed by ISPI. The question of analysis method then has a simple response: we analyze both the ISPI microscope image and the resist moiré image using the same algorithm.

With the above considerations in mind, we designed a special alignment test mask and conducted the x-ray exposure tests as follows. We wrote on a single x-ray mask that contained 96 of the $\{p_1, p_2\}$ alignment marks appropriate for ISPI, plus a variety of additional patterns. The alignment mark gratings had periods of 1.00 and 1.05 μm. A first x-ray exposure in contact transferred the mask patterns onto a silicon substrate,

where they were subsequently etched in relief using reactive-ion etching. The substrate was then recoated with SAL 601 resist and placed back in the ISPI alignment apparatus. The x-ray mask was rotated 180 deg. which, by virtue of its designed symmetry, brought $\{p_1, p_2\}$ gratings on the mask into superposition with $\{p_2, p_1\}$ marks on the substrate (i.e., a 180 deg. rotation converted $\{p_1, p_2\}$ pairs to $\{p_2, p_1\}$ pairs). The mask was aligned and lowered onto the substrate, where it went into intimate contact. A lateral shift of ~100 nm occurred when the mask went into contact[30]. After the mask was in intimate contact, we observed and recorded the phase disparity between the two halves of each mark, and converted that into a displacement, called ΔX_1. Following exposure, we developed the resist and measured the overlay of the $\{p_1, p_2\}$ resist patterns on the $\{p_2, p_1\}$ relief gratings in Si for all of the mark pairs, using a normal-incidence Leitz microscope. A representative example of both fringe pairs is shown in Fig. 9.1.

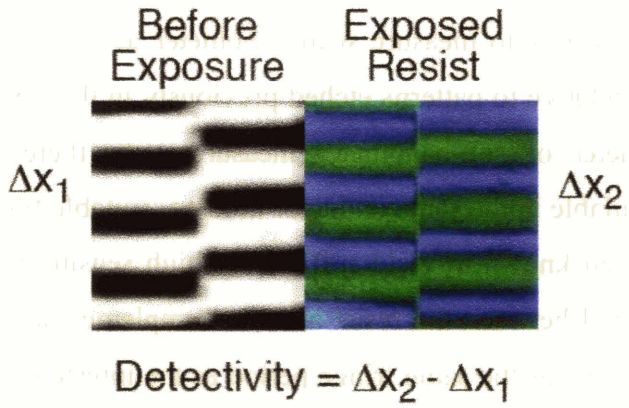


Fig. 9.1. (a) Fringe pattern observed through ISPI microscope. (b) Fringe pattern in a moiré pattern between exposed resist and grating in substrate, observed by a Leitz microscope.

The measured displacement of the moiré image of resist on a silicon relief grating is called ΔX_2. We call the quantity $(\Delta X_1 - \Delta X_2)$ the "disparity between detected and

[30] The lateral shift is attributed to small inequalities in the height of the mesa surrounding the mask membrane. The shift occurred despite leveling the mask prior to bringing it into contact with the substrate, and varied by 10s of nm between repeated contact cycles.

measured misalignment". The alignment difference was found for 24 marks. The difference of these measurements is plotted in Fig. 9.2 for marks with 1.00 and 1.05 μm grating period. Alignment as perceived by ISPI and alignment shown in the exposed pattern agreed, with a mean difference of $\delta = 0.1$ nm, with $\sigma = 0.9$ nm.

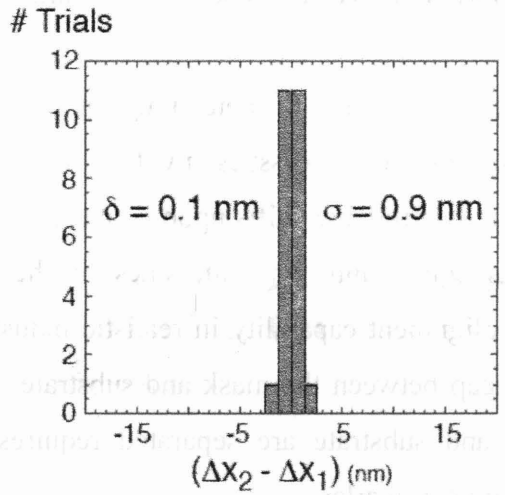


Fig. 9.2. ISPI alignment detectivity, showing a histogram of $(\Delta X_1 - \Delta X_2)$ for the 1 μm-period alignment marks. The difference of moiré fringes as detected in an ISPI microscope before exposure and fringes observed in resist with a different microscope at normal incidence after exposure and development. The x-ray exposure is done with the mask in contact with the substrate to eliminate the effect of point-source magnification. The mean difference was 0.1 nm and $\sigma = 0.9$ nm.

We believe this experiment proves that there is a correspondence between the misalignment, as *detected* by the ISPI optics and algorithm, and the misalignment as *measured* in x-ray-exposed resist. We believe the results demonstrate that the ISPI scheme has a detectivity of misalignment better than 1 nm, with a standard deviation of <1 nm.

It should be emphasized that the same marks are used for both measurements, yet the means of detection are quite distinct. There is a substantial difference between the

two types of fringes, in both their underlying physical formation and the observation mechanisms. The ISPI fringes are viewed at an angle of 20° using the low-NA, low magnification ISPI microscopes with multiple laser line illumination, while the fringes in resist are viewed at normal incidence with a high-NA, high magnification microscope and white light. Despite these differences in the means of viewing the fringes, the disparities between the two forms of measurements were <1 nm.

This experiment gives a measure of detectivity that is independent of many undesirable influences. However, due to issues involved with mask-substrate contact (risk of breaking the mask, uncontrollable shifts upon contact, particulate contamination, time required to reach gas equilibrium on both sides of the membrane, etc.), this experiment does not prove alignment capability in realistic industrial conditions, which involve a small, but finite, gap between the mask and substrate. Nanometer-level drift occurring when the mask and substrate are separated requires continuous feedback stabilization to achieve nanometer overlay.

9.2 Overlayed Exposures with Feedback Control

In the following sections we will examine open-loop behavior, stability of the ISPI alignment system when using feedback-controlled positioning, and then demonstrate 72-hour aligned exposures[31] that define the upper and lower gates of a double-gate transistor.

9.2.1 Drift and Vibration

Figure 9.3 shows an example of the drift in alignment detected during six hours at a mask-substrate gap of 50 microns. The piezo drives were held stationary, and the output

[31] Such exposure times are not viable in industrial use, but exemplify the capabilities of feedback control.

of a single ISPI microscope was monitored for six hours. The relative mask/substrate alignment drifted slowly through about 150 nm, with faster, smaller fluctuations along the way. This data was taken during the night to minimize disturbances from laboratory activities.

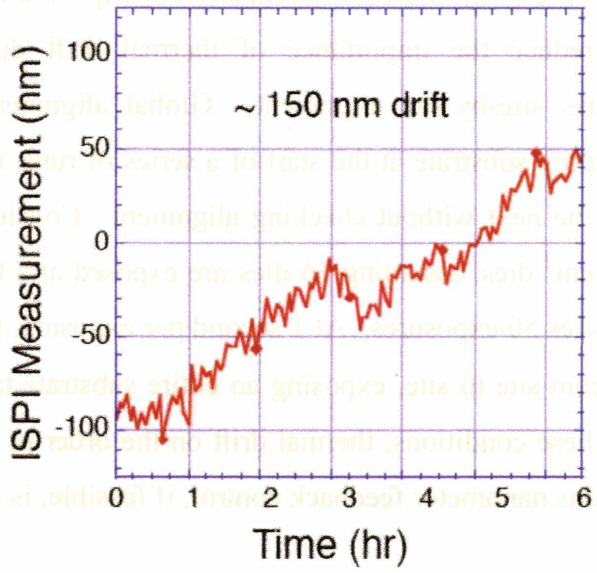


Fig. 9.3. 150 nm drift in alignment over six hours.

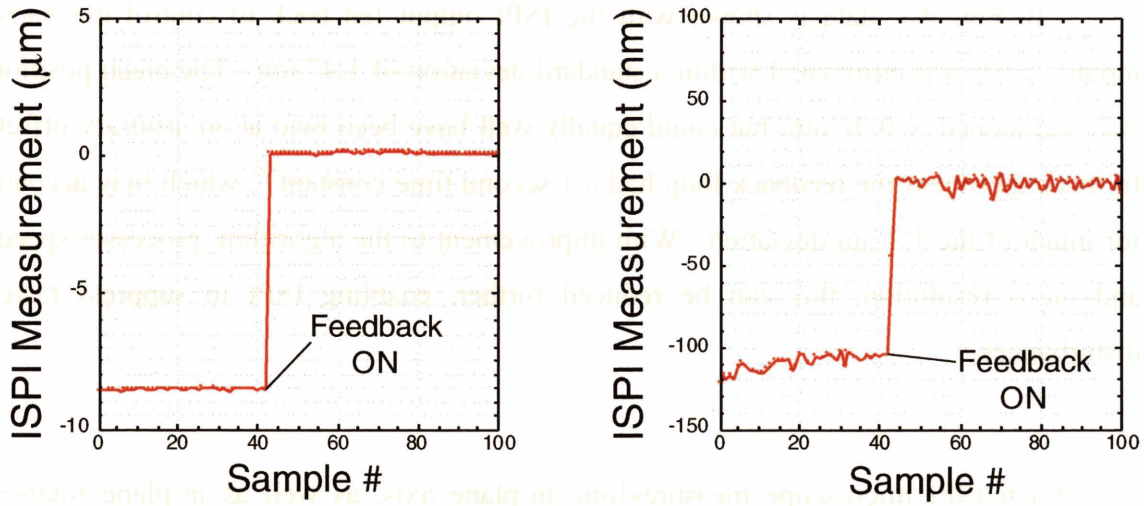
One solution to thermal drift, which has been employed at MIT, is to place the mask and substrate in mechanical contact, e.g., set the mask on Al studs of a few microns. Friction holds the mask and substrate together. Alignment must be done when the mask studs are above the substrate. Just before exposure, the mask is lowered a few microns to rest on the studs. Lowering the mask inevitably introduces an unpredictable misalignment of 10s of nanometers. When the studs are in contact with the substrate the alignment is *stable* to sub-nanometer levels, but the unpredictable misalignment inherent in this procedure is incapable of producing nanometer-level overlay.

Unlike our laboratory, in an industrial setting the environment of a lithographic tool is temperature-controlled to <0.1 C. This certainly will reduce thermal drift, but on

the nanometer scale, a temperature change of only 0.1° C can still result in drift of tens of nanometers. Another common practice is to construct the lithographic tool on a granite superstructure to improve structural rigidity and reduce vibrations. Finally, shorter exposure times in industry place less stringent requirements on long-term stability. If a synchrotron is used as the x-ray source, the reduction in exposure time (~1 sec exposure duration) will also reduce the importance of thermal drift during each exposure. However, this assumes site-by-site alignment. Global alignment, which is used in industry to align an entire substrate at the start of a series of runs, relies on the stages to step from one die to the next without checking alignment. Considering that a 300 mm substrate has ~800 1-cm^2 dies, assuming 16 dies are exposed at a time, this implies that the full substrate requires 50 exposures. At 1 second per exposure, plus an overhead time of 1 second to step from site to site, exposing an entire substrate takes 50x2s = 100 s = 1.7 minutes. Under these conditions, thermal drift on the order of tens of nanometers is possible, and continuous nanometer feedback control, if feasible, is a desirable prospect.

9.2.2 Feedback Control

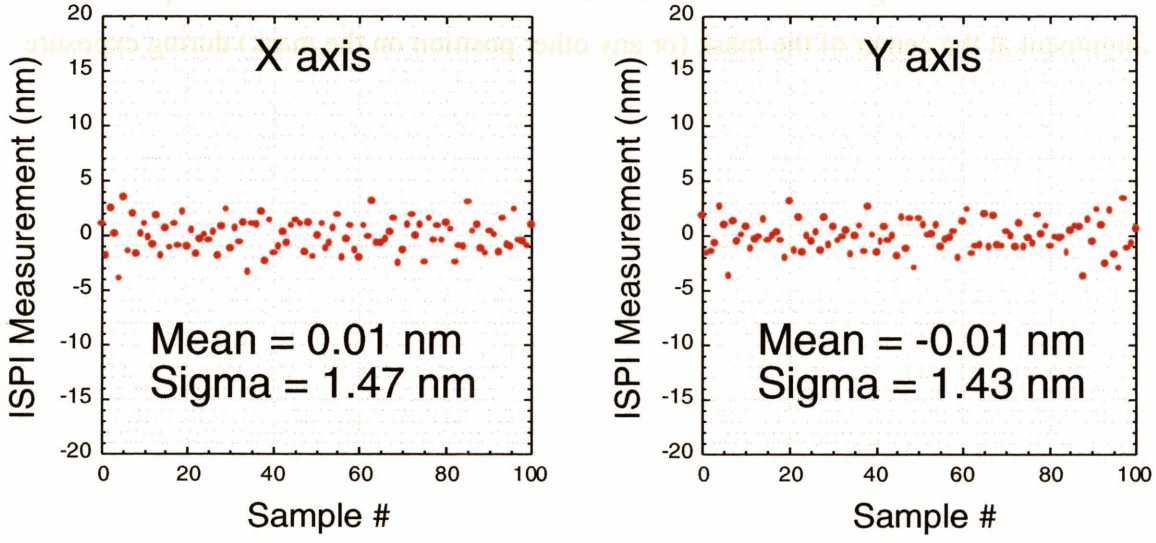
The alignment feedback loop consists of the alignment fringes imaged by the microscopes onto CCD cameras, image digitization by a framegrabber, extraction of the alignment condition by image analysis software, coordination of position measurements from multiple axes, and appropriate piezo actuation. Depending on the amount of misalignment, correction signals are sent either to the substrate stage or to the piezos in the mask stage. For misalignments larger than 250 nm, the substrate stage makes corrections, but smaller corrections are made by the piezo-actuated mask stage. Figure 9.4 shows data taken during alignment acquisition, starting from a large misalignment.



(a) (b)

Fig. 9.4. Plots of the response of the ISPI feedback system to large misalignments. (a) Data from a bar array. An initial displacement of ~9 µm is corrected using the spatial phase disparity between mask and substrate bar arrays. (b) Data taken simultaneously from ISPI fringes. The residual misalignment is corrected using the fringe-fringe phase disparity. One sample step was taken every second.

The efficacy of feedback in maintaining position is illustrated by the experimental results depicted in Fig. 9.5.



(a) (b)

Fig. 9.5. Plots of feedback control in (a) X axis and (b) Y axis. The sample rate is 1 Hz.

In Fig. 9.5, data is shown with the ISPI output fed back to control the piezo inputs. Drift was eliminated within a standard deviation of 1.47 nm. The mean position was maintained at 0.01 nm, but could equally well have been held at an arbitrary offset. In this experiment the feedback loop had a 1 second time constant[32], which may account for much of the 1.4 nm deviation. With improvement to the algorithm, processor speed, and piezo resolution, this can be reduced further, enabling ISPI to suppress faster disturbances.

Each ISPI microscope measures one in-plane axis, as well as in-plane rotation (indicated by a fringe chevron pattern) and gap. Three ISPI microscopes are used for detection of all six spatial degrees of freedom: X, Y, Z, θ_x, θ_y, θ_z. One microscope measures X and two measure Y at a finite separation. X and Y microscopes can measure alignment in the same area with adjacent X and Y marks. Rotation is derived most sensitively from the difference in the two Y measurements. Gap is detected by each microscope using TCG marks at exposure gaps, yielding Z_1, Z_2, Z_3, from which mask rotations are derived, completing the measurement of all six degrees of freedom.

A Littrow angle of 20° is sufficient to allow the ISPI microscopes to detect alignment at the center of the mask (or any other position on the mask) during exposure.

[32] These feedback experiments were done using a 1.8 GHz Pentium 4.

9.2.3 Overlayed Exposures

To verify the accuracy of ISPI alignment, we utilize a point x-ray source to expose "mask" alignment marks on top of "substrate" alignment marks. The magnification effect of the divergent x-ray source causes a lateral shift of the projected mark image, which increases with radial distance from the source axis. To measure overlay with minimum magnification effect, we position an alignment mark on the source axis prior to exposure of resist (by observing the divergent x-rays through a micropipette bundle on an x-ray-sensitive CCD, as in Section 8.5), and measure the overlay in two successive exposures. Phase analysis of the moiré between the resist gratings on top of the gratings in the substrate permits sub-1-nm overlay detection. This approach includes all error terms intrinsic to the alignment system.

In this experiment we exposed a mask with "mask" marks and "substrate" marks onto a silicon substrate using UV-5 resist, as illustrated in Fig. 9.6.

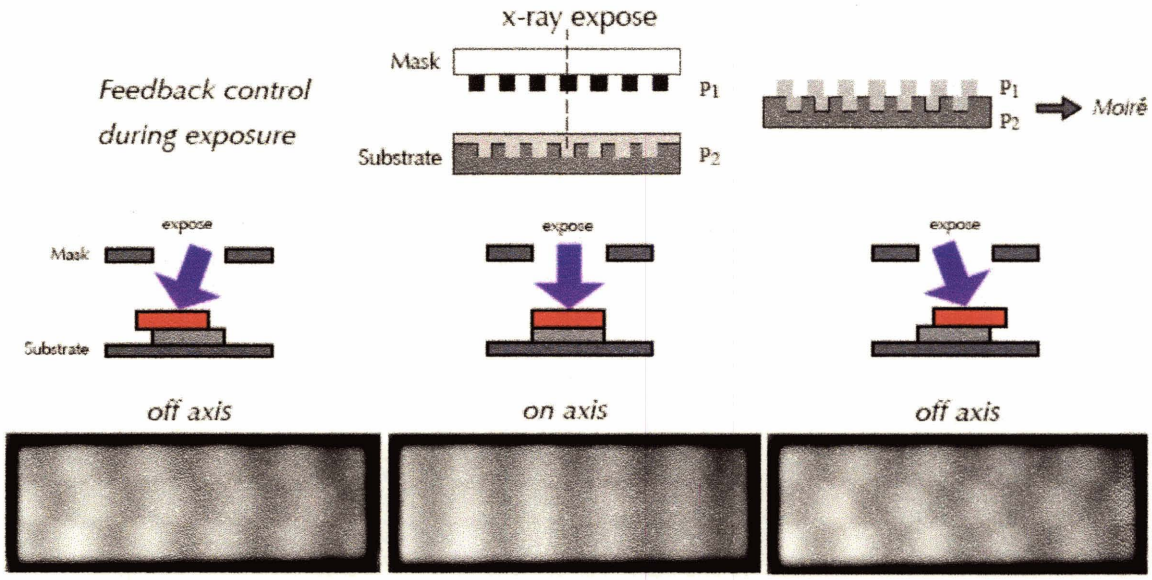


Fig. 9.6. Schematic of a feedback-controlled overlay experiment. Alignment was maintained in X, Y and θ throughout the exposure. Images of the moiré in exposed marks indicated the center mark was aligned to the source axis. Overlay at the mark under the source axis was 2.7 nm.

The "mask" mark is composed of three gratings with periods, $\{p_1, p_2, p_1\}$ and the "substrate" mark consists of these same grating periods, but with the sequence interchanged $\{p_2, p_1, p_2\}$. These $\{p_2, p_1, p_2\}$ marks were used for detection of the alignment in resist. The marks viewed by the ISPI microscopes during alignment and exposure were adjacent to these marks, and of the standard grating-over-checkerboard type. Both marks were written within the same e-beam field to avoid stitching errors. The marks were printed at a 3 μm gap set by aluminum studs. After the first exposure the marks were etched into the silicon. The substrate was recoated with UV-5 and the $\{p_2, p_1, p_2\}$ mark aligned underneath the $\{p_1, p_2, p_1\}$ mark on the mask. Alignment and rotation of the mask were controlled actively throughout the exposure. The resist was developed and the moiré from superposition of (i) gratings in resist and (ii) gratings in relief in silicon was observed in a normal-incidence optical microscope and analyzed with the ISPI spatial phase algorithm described in Section 3.5. The detected overlay was 2.7 nm. This overlay error combines alignment detection errors, source-positioning errors, intrafield e-beam writing errors, grating imperfections, as well as other contributions within the system.

9.2.4 *Aligned Double-Gate Transistor Structure*

In collaboration with another graduate student[33], device structures for a double-gate MOSFET transistor were fabricated using the ISPI alignment system. Such a double-gate device has superior performance characteristics and low power consumption, and is considered to be an important development for transistor scaling below 50 nm gate widths. The challenge in fabrication of such a device is the superposition of the two gates. Self-alignment avoids this problem, but raises other significant fabrication difficulties. In this project, we took the conceptually simpler, more direct approach of alignment control of both upper and lower gates. A schematic of a double-gate MOSFET is shown in Fig. 9.7. Figure 9.8 shows feedback data during an aligned exposure, and Fig. 9.9 shows optical micrographs of the fringes exposed in resist, and the macroscopic device structure. Figure 9.10 shows a SEM cross-section of the resulting gate structures.

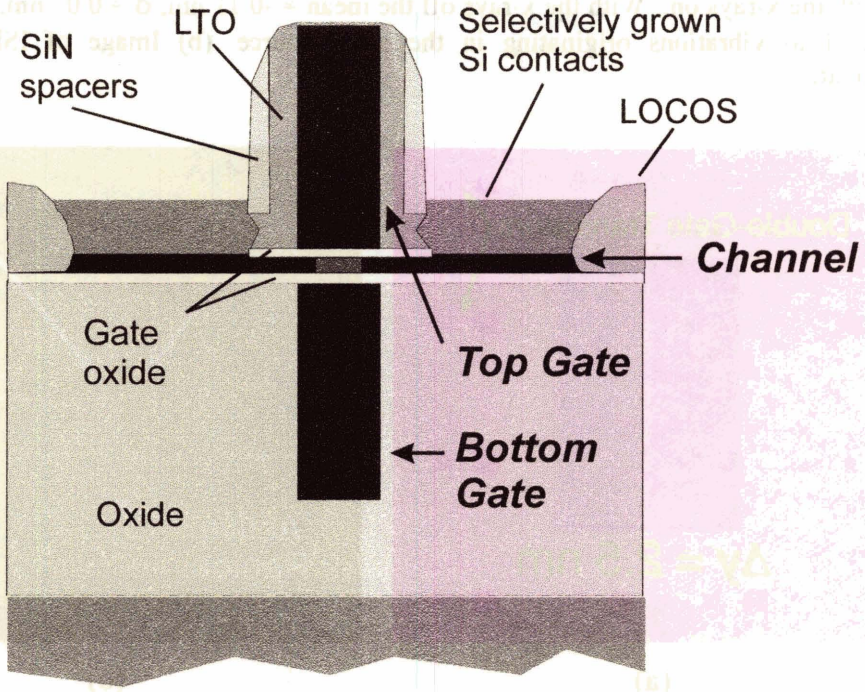
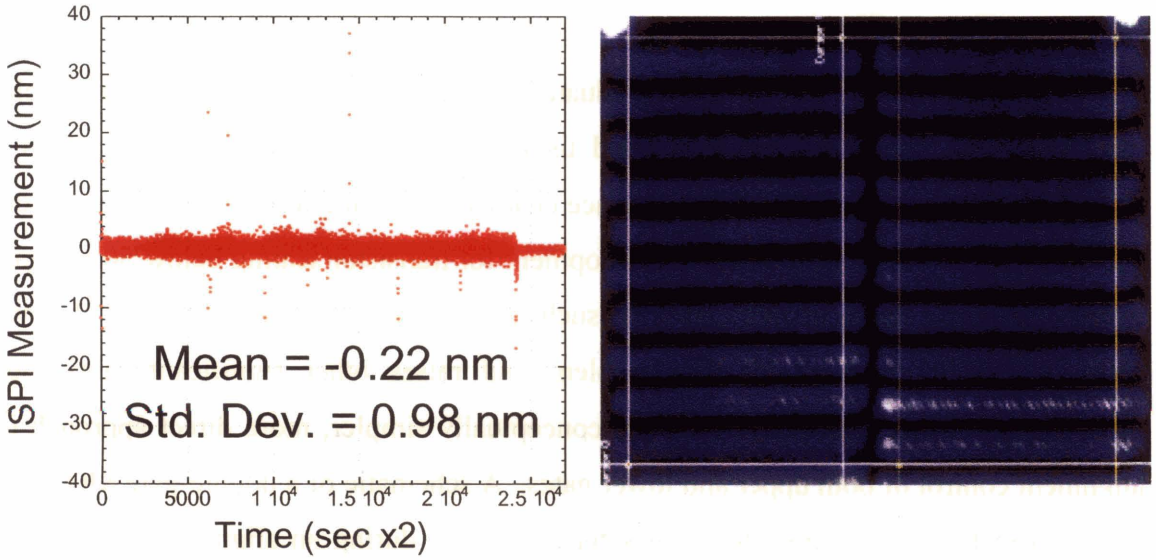


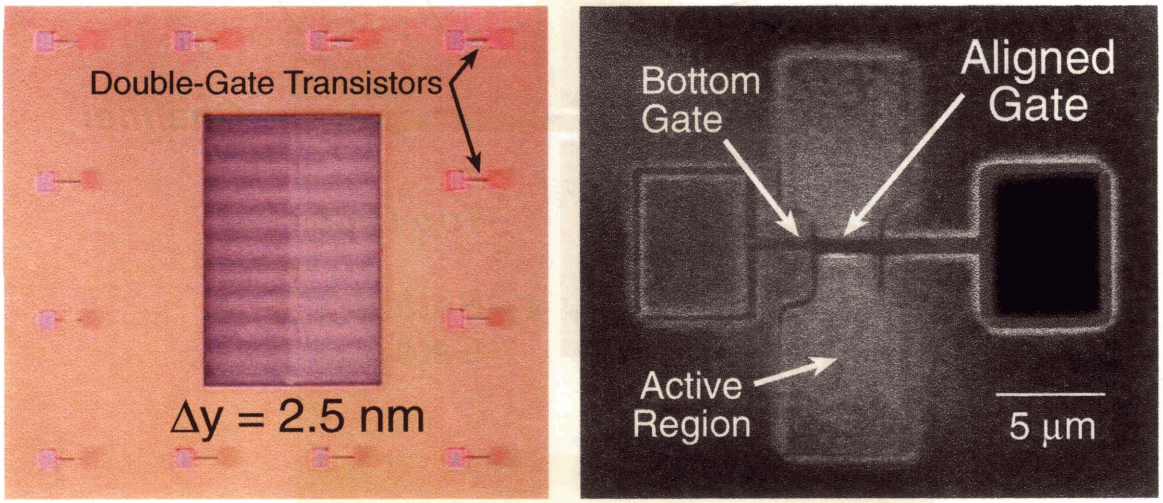
Fig. 9.7. Schematic of a double-gate MOSFET device. (Courtesy of Dr. Meinhold.)

[33] Dr. Mitchell W. Meinhold



(a) (b)

Fig. 9.8. (a) Y-alignment data from an ISPI microscope viewing a mark adjacent to a double-gate transistor. The mean alignment error detected by the ISPI microscope was -0.22 nm and $\sigma = 0.98$ nm, with the x-rays on. With the x-rays off the mean = -0.11 nm, $\sigma = 0.07$ nm. The difference is attributed to vibrations originating in the x-ray source. (b) Image of ISPI fringes during alignment.



(a) (b)

Fig. 9.9. (a) Optical micrograph of gates and fringes in resist after exposure. The alignment error detected by the fringes in resist was measured to be 2.5 nm. (b) Optical micrograph of double-gate structure.

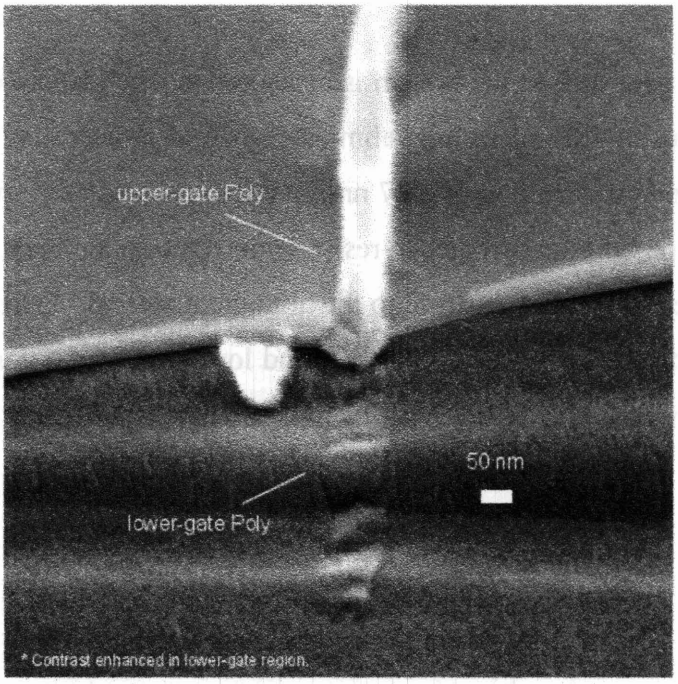


Fig. 9.10. SEM micrograph of 50 nm gate structures fabricated using ISPI feedback-controlled overlay. Fringes in resist adjacent to the gates indicate misalignment to be 2.5 nm. (Image courtesy of Dr. Meinhold.)

Further details of the double-gate MOSFET work can be found in Dr. Meinhold's PhD thesis [9.10].

9.3 Summary

This chapter contains the lithographic overlay experiments that verify the ISPI alignment capability. Section 9.1 demonstrates the detectivity of ISPI, i.e., the difference between the spatial phase of the interference fringes seen in the ISPI microscopes, and the moiré fringes between exposed resist and etched silicon on the substrate. The exposure is performed with the mask in contact with the substrate, to avoid all possible deviations associated with mechanical drift during the exposure, as well as magnification errors related to the x-ray point source.

At a finite gap, drift in the system is measured to be >150 nm over several hours. Feedback control using ISPI metrology eliminates drift, demonstrating a mean position of 0.01 nm and a standard deviation of 1.47 nm. After exposures performed with feedback, examination of exposed moiré fringes in resist demonstrates alignment overlay to 2.7 nm. This includes all error sources intrinsic to the alignment system. During consecutive 72-hour feedback-controlled exposures the upper and lower gates of a double-gate transistor structure are demonstrated to be overlayed within 2.5 nm.

10 Conclusions

The goal in this dissertation is to look beyond the alignment needs of a single type of lithography, or a single generation of fabrication technology. Instead, the purpose is to look towards a broadly applicable alignment technology that can extend to the limits of fabrication, and improve position detection to the point that it is no longer an impediment to any presently conceivable form of proximity fabrication. Toward this end, in this dissertation we developed a novel six-axis nanopositioning scheme called Interferometric-Spatial-Phase Imaging (ISPI) that provides robust sub-nanometer detectivity in measurement of alignment, rotation, and gap through spatial phase analysis of interferometric moiré fringe patterns. Littrow-angle imaging of the fringes enables continuous feedback control of all six axes, and employs diffractive dark-field microscopy, increasing signal-to-noise ratio in the image. Experimental evidence of nanometer overlay of exposed features, in periodic resist patterns and double-gate field-

effect transistor structures, indicates ISPI can meet the nanopositioning requirements in the projected future.

10.1 Advantages of ISPI

A concise summary of some of the advantages of ISPI:

1) Moiré fringes have arbitrarily high sensitivity. Sensitivity is limited by the microscope field of view and the grating periods. The fact that we use interference to form the moiré pattern does not restrict this to interfering beams, but the same kind of scheme would work well with obstruction or sampling-based moiré [10.1]. Thus the same principles will apply to a wide variety of alignment configurations.
2) Moiré can be evaluated in the frequency domain, which allows one to utilize the advantages of phase measurements and frequency domain analysis.
3) Arbitrary positions, not just the zero, or null position, can be measured with equal accuracy.
4) ISPI is wavelength-independent. Broadband light is used, eliminating thin-film interference effects. Laser requirements are relaxed since the laser wavelengths are free to drift without any deleterious effect on the spatial phase.
5) Displacement measurement does not demand a displacement of the mask or substrate, i.e., no scanning is required.
6) The measurement is absolute and unambiguous. Since ISPI is based on imaging, multiple periodic marks can be viewed in a single image. We exploit the difference in displacement magnification between geometric motion and interferometric moiré to obtain different displacement sensitivity. The marks are designed with the same period allowing the same spatial phase algorithm to be used to find the displacement of the individual marks, and the phase between

marks, providing a unique and unambiguous position measurement. The acquisition range, i.e., the maximum misalignment, is at least half the largest period, or half the fringe period. In practice we have shown wide-rage, unambiguous measurement, from <0.5 nm to >500 μm.

7) Imaging not only provides absolute and unambiguous measurement, it also permits measurement along multiple axes in the same image. Imaging with ISPI permits alignment detection in the direction transverse to the plane containing the viewing angle, and also allows detection of rotation, by virtue of the relative angle between the complementary fringe sets (i.e., chevron pattern), as well as simultaneous gap detection.
8) Since each ISPI microscope views alignment (and rotation and gap) at an oblique angle, we do not need to artificially separate alignment from lithography: the two can proceed simultaneously, and if any misalignment occurs during the exposure, we can correct for it immediately.
9) ISPI extracts pertinent information since it measures alignment between the mask and substrate, as opposed to measuring the relative position of a pair of stages, as in a heterodyne interferometer.
10) The ability to detect alignment during exposure becomes more of a necessity as the exposure time increases, and as the alignment tolerance decreases to the point that thermal fluctuations are comparable or in excess of the alignment tolerance. When this is the case, we can still perform highly accurate alignment even without expensive and cumbersome environmental controls, or construction of the stages or other parts of the stepper with zero-thermal-expansion materials.
11) Noise reduction: there are four stages of filtering that ISPI employs to decrease spurious signals. First, it eliminates the zero-order beams from the image formation – they reflect off the mask and substrate, out of the acceptance angle of the low-NA microscope. Second, the microscope is diffractive dark-field, that is, the only light that returns to contribute to the image is signal-carrying light that was designed to diffract and return at a specific angle. No other light makes it into the microscope except the light that contains the spatial phase information

(with the exception of a negligible amount of scattered light from the miniscule area of feature edges facing the microscope). Third, since the microscope cannot resolve high spatial frequencies; it acts as a low pass filter. We select the resolution of the microscope to cut off all spatial frequencies above the moiré fringe frequency. The microscope therefore cannot resolve the grating lines, dust, defects, or pattern features that commonly corrupt the purity of the signal. Furthermore, after the lowpass filtering by the microscope, the phase analysis algorithm performs narrow bandpass filtering of the remaining signal to select only the known spatial frequency of the moiré fringes.

12) Any systematic bias of one set of fringes is exactly balanced by corresponding changes in the complementary set of fringes. It is the phase relationship between these fringes that determines the alignment signal; therefore the signal is immune to common perturbations of interferometer signals, such as variations in the index of refraction in the beampath. Both signals are affected equally, unless in the rare case that high turbulence over very short length scales (diameter of the beam, i.e., a few hundred microns) causes disparities between the two sides.
13) Motion of the camera does not affect the relative spatial phase of the fringe sets.
14) Transverse Chirp Gapping (TCG) detects gap with <1 nm sensitivity. No marks are required on the substrate. TCG is independent of alignment.
15) Long-Range Interferometric Gapping (LRIG) extends the <1 nm gap sensitivity of TCG up to gaps of 500 mm, or more.
16) Most importantly, the ISPI philosophy is to measure the variable of interest, at the point of importance, and by such means that the measurement signal is impervious to common environmental degradations.

ISPI is unique in that it combines interferometric sensitivity, with immunity to errors of spatial phase, while providing static measurement, noise reduction, and multi-axis imaging capability.

10.2 Resolution Improvement

Simple improvements to alignment detectivity could be achieved by reducing the grating period, increasing the number of camera pixels and reducing camera signal to noise ratio.

Resolution improvements are clearly possible, and should be relatively simple to implement. Considering Eq. 3.36, the best way to improve resolution is to decrease the grating period. We typically use p_1 and p_2 gratings of ~1 μm period, with non-exposing wavelengths between 655 and 690 nm. We have successfully used periods as small as 0.8 microns with these laser wavelengths. Slightly shorter wavelengths could be used, in fact we have used wavelengths in the green (515 nm) and blue (487 nm) for alignment, but of course the resist sensitivity at those wavelengths must be considered if the alignment marks are near devices. Since the ISPI grating configuration is achromatic, much shorter wavelengths are possible. A 405 nm illumination wavelength with a 0.5 μm grating period would double the phase estimation accuracy.

Other factors are SNR and number of samples. In the overlay experiments in Chapter 9 we used analog CCD cameras with VGA (640x480) pixel arrays. Commercial digital machine vision cameras are available with 4 megapixel arrays, increasing pixel count to (2350x1720), and increasing phase estimation accuracy by an additional factor of $(3.69)^{0.5} = 1.9$. Experiments on digital CMOS cameras indicate a ~20% reduction in pixel noise compared to the analog cameras. Averaging along more pixels in the fringe direction will yield another factor of 1.9. Taking the simple modifications of the alignment mark periods and replacement of the cameras and laser source into account, phase estimation accuracy would be improved by ~7.2x. In short, the resolution of the system could be improved to ~0.010 nm, with minimal effort.

11 References

[1.1] www.shef.ac.uk/uni/projects/hp/microms4.htm

[1.2] Metrology2003-1.pdf (public.itrs.net/Files/2003ITRS/Home2003.htm)

[1.3] Ahmed Busnaina, Director of the Nanomanufacturing Research Institute (www.masshightech.com)

[1.4] J. Cheng, Fabrication and Characterization of Nanostructures from Self-Assembled Block Copolymers, Ph.D. Thesis, MIT, Cambridge, MA (June 2003).

[1.5] THOSE WHO CAME BEFORE, Southwestern Archeology in the National Park System, by Robert H. Lister and Florence C. Lister (1993).

[1.6] HOHOKAM INDIANS OF THE TUCSON BASIN, by Linda M. Gregonis and Karl J. Reinhard (1979).

[1.7] HP Technical Description and Specifications, Laser Interferometer HP 5525A (June 1, 1970).

[2.1] H.I. Smith, D.L. Spears, and S.E. Bernacki, "X-ray Lithography: A Complementary Technique to Electron Beam Lithography", J. Vac. Sci. Technol. 10(6), 913 (1973).

[2.2] K.R. Early, Experimental Characterization and Physical Modeling of Resolution Limits in Proximity Printing X-ray Lithography, Ph.D. Thesis, MIT, Cambridge, MA (June 1991).

[2.3] Suss MicroTec Inc., 228 Suss Drive, Waterbury Center, VT 05677 (www.suss.com)

[2.4] A. Flamholz and R. Rippstein, J. Vac. Sci. B **8** 2002 (1990).

[2.5] E. Cullmann and K. Cooper, J. Vac. Sci. Technol. B **6**, 2132 (1988).

[2.6] T. Miyatake, M. Hirose, T. Shoki, R. Ohkubo, and K. Yamazaki, Proc. SPIE 3048, 225 (1997).

[2.7] T. Miyatake, M. Hirose, T. Shoki, R. Ohkubo, and K. Yamazaki, J. Vac. Sci. Technol. B **15**, 2471 (1997).

[2.8] T. Miyatake, M. Hirose, T. Shoki, R. Ohkubo, and K. Yamazaki, J. Vac.

Sci. Technol. B **16**, 3471 (1998).

[2.9] T. Miyatake, X. Li, S. Hirose, T. Monzen, K. Fujii, and K. Suzuki, "Compact Synchrotron Radiation Lithography System for 70 nm Device Manufacturing", J. Vac. Sci. Technol. B **19**, 2444 (2001).

[2.10] B. Fay, J. Trotel, and A. Frichet, J. Vac. Sci. Technol. **16**, 1954 (1979).

[2.11] G. Chen, F. Cerrina, and Z. Wu, "X-ray Lithography Two-State Alignment System", J. Vac. Sci. Technol. B **7**, 1995 (1989).

[2.12] G. Chen, J. Wallace, F. Cerrina, S. Palmer, B. Newell, and J. Randall, J. Vac. Sci. Technol. B **9**, 3222 (1991).

[2.13] G. Chen, J. Wallace, R Hachman, G. Wells, D Bodoh, P. Anderson, M. Reilly, and F. Cerrina, "CxrL Aligner: An Experimental X-ray Lithography System for Quarter-Micron Devices", J. Vac. Sci. Technol. B **10**, 3229 (1992).

[2.14] H. Zhou and M. Feldman, "Sub-nanometer Alignment System for X-ray Lithography", J. Vac. Sci. Technol. B **12**, 3261 (1994).

[2.15] D.C. Flanders H.I. Smith, and S. Austin, Appl. Phys. Lett. 31, 426 (1977).

[2.16] S. Austin, H.I. Smith, and D.C. Flanders, "Alignment of X-ray Lithography Masks Using a New Interferometric Technique - Experimental Results ", J. Vac. Sci. Technol. 15, 984 (1978).

[2.17] M. Nelson, J.L. Kreuzer, and G. Gallatin, "Design and Test of a Through-The-Mask Alignment Sensor for a Vertical Stage X-ray Aligner ", J. Vac. Sci. Technol. B **12**, 3251 (1994).

[2.18] M. Suzuki and A. Une, , "An optical-heterodyne alignment technique for quarter-micron x-ray lithography ", J. Vac. Sci. Technol. B **7**, 1971 (1989).

[2.19] J. Itoh, T. Kanayama, N. Atoda, and K. Hoh, "An alignment system for synchrotron radiation x-ray lithography", J. Vac. Sci. Technol. B **6**, 409 (1988).

[2.20] M. Tabata and T. Tojo, "High-Precision Interferometric Alignment Using Checker Grating", J. Vac. Sci. Technol. B **7**, 1980 (1989).

[2.21] K. Hara, Y. Uchida, T. Nomura, S. Kimura, D. Sugimoto, A Yoshida, H. Miyake, and T. Iida, "An Alignment Technique Using Diffracted Moiré Signals ", J. Vac. Sci. Technol. **7**, 1977 (1989).

[2.22] Lord Raleigh, "On the manufacture and theory of diffraction gratings", Phil. Mag. Ser. 4, 47(310), 81-93 (February 1874) and 47(311), 193-205 (March 1874).

[2.23] Lord Raleigh, "On copying diffraction-gratings, and on some phenomena connected therewith", Phil. Mag. Ser. 11, 196-205 (1881).

[2.24] J. Guild, The Interference Systems of Crossed Diffraction Gratings - Theory of Moiré Fringes, Oxford University Press, London, England, 1956.

[2.25] J. Guild, Diffraction Gratings as Measuring Scales, Oxford University Press, London, England, 1960.

[2.26] P. Konkola, Design and analysis of a scanning beam interference lithography system for patterning gratings with nanometer-level distortions, Ph.D. Thesis, MIT, Cambridge, MA (June 2003).

[2.27] C. Chen, Beam Alignment and Image Metrology for Scanning Beam Interference Lithography – Fabricating Gratings with Nanometer Phase Accuracy, Ph.D. Thesis, MIT, Cambridge, MA (June 2003).

[2.28] M.C. King and D.H. Berry, "Photolithographic Mask Alignment Using Moire Techniques," Applied Optics, Vol 11(11), 2455 (1972).

[2.29] J.G. Goodberlet, B.L. Dunn, Deep-Ultraviolet Contact Photolithography, Micro & Nano Engineering, 1999, 11-12. (1999).

[2.30] T.M. Lyszczarz D.C. Flanders, N.P. Economou, and P.D. DeGraff, "Experimental evaluation of interferometric alignment techniques for multiple mask registration ", J. Vac. Sci. Technol. 19(4), 1214 (1981).

[2.31] A. Moel, E. E. Moon, R. Frankel, and H. I. Smith, J. Vac. Sci. Technol. B **11**, 2191 (1993).

[3.1] A. S. Willsky, G. W. Wornell, and J. H. Shapiro, Stochastic Processes Detection and Estimation, Cambridge, MA, Massachusetts Institute of Technology, 1998.

[4.1] www.physikinstrumente.de/products/prdetail.php?secid=5-4

[4.2] D. C. Flanders and T.M. Lyszczarz, "A Precision Wide-Range Optical Gap Measuring Technique ", J. Vac. Sci. Technol. B 1, 1196 (1983).

[4.3] M. Suzuki and A. Une, , "An optical-heterodyne alignment technique for quarter-micron x-ray lithography ", J. Vac. Sci. Technol. B 7, 1971 (1989).

[6.1] H.F. Talbot, *Phil. Mag.* **9** (1836) 401-407.

[9.1] H. I. Smith, S. D. Hector, M. L. Schattenburg, and E. H. Anderson, J. Vac. Sci. Technol. B **9**, 2992 (1991).